

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 18:44:25 ; Search time 4324 Seconds
(without alignments)
4630.604 Million cell updates/sec

Title: US-09-700-590A-22

Perfect score: 3639
Sequence: 1 MSAAEGPGTRRLNLPVMDG.....ALCHLLFAAEKETFRLRLM 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=1p
-O=/cgn2.1/USPTO_spool/US9700590/runat_14032003_140630_3224/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09700590@cgn 1 1 3745 @runat_14032003_140630_3224 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERIES -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGABEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb_ba :
2: gb_hcg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :

29: em_vi :
30: em_hcg_hum :
31: em_hcg_inv :
32: em_hcg_other :
33: em_hcg_mus :
34: em_hcg_pln :
35: em_hcg_rtd :
36: em_hcg_mam :
37: em_hcg_vrt :
38: em_sy :
39: em_htgo_hum :
40: em_htgo_mus :
41: em_htgo_other :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	3583	98.5	3149	9 HSHUNKI	Y12059 H.sapiens H
2	3583	98.5	5198	9 AF386649	AF386649 Homo sapi
3	3441.5	94.6	2727	10 AF461396	AF461396 Mus muscu
4	3436.5	94.4	5281	10 AF273217	AF273217 Mus muscu
5	3436.5	94.4	6000	10 AF461395	AF461395 Mus muscu
6	2566.5	70.3	2199	10 BC008532	BC008532 Mus muscu
7	1931	53.1	3033	10 BC031536	BC031536 Mus muscu
8	1928	53.0	2933	10 AF269193	AF269193 Mus muscu
9	1922	52.8	3028	9 HUMOREX	D26362 Human mRNA
10	1901	52.2	3649	10 AF045462	AF045462 Mus muscu
11	1888	51.9	4053	9 HSPSHMR	X62083 H.sapiens m
12	1888	51.9	4053	9 HUMFSG	M80613 Human homol
13	1888	51.9	4664	6 AA409010	AA409010 Sequence
14	1888	51.9	4664	6 HMKIA9001	D42040 Human mRNA
15	1877.5	51.6	2548	9 BC032124	BC032124 Homo sapi
16	1837	50.5	4667	9 HSM804033	AF183272 Homo sapi
17	1684	46.3	2522	5 AF191032	AF191032 Myxine gl
18	1573	43.2	37644	6 A85302	A95302 Sequence 33
19	1573	43.2	92863	12 GGBLOCUS	AL023516 Gallus ga
20	1548.5	42.6	4651	5 GGRING3GE	X96669 G.gallus RI
21	1540.5	42.3	3429	9 BC011898	BC011898 Homo sapi
22	1540	42.3	3391	10 AF358660	AF358660 Mus muscu
23	1535.5	42.2	2350	9 BC000477	BC000477 Homo sapi
24	1520	41.8	1657	10 AB010247	AB010247 Mus muscu
25	1520	41.8	1688	10 AB010246	AB010246 Mus muscu
26	1501.5	41.3	3106	9 AF019085	AF019085 Homo sapi
27	1500	41.2	1554	10 AB010248	AB010248 Mus muscu
28	1458.5	40.1	3404	9 AK056504	AK056504 Homo sapi
29	1451.5	39.9	11493	10 MUI1KLT	AL009226 Mus muscu
30	1451.5	39.9	273800	10 AF100956	AF100956 Mus muscu
31	1450	39.8	1884	9 AK024312	AK024312 Homo sapi
32	1441	39.6	10222	10 D89801	D89801 Mus muscu
33	1441	39.6	216734	5 AB073376	AB073376 Oryzias l
34	1440.5	39.6	274856	2 AC098547	AC098547 Rattus no
35	1401	38.5	2206	10 AF318183	AF318183 Mus muscu
36	1380.5	37.9	5632	3 DROFSHB	M23222 D.melanogas
37	1380.5	37.9	7792	3 DROFSHA	M23221 D.melanogas
38	1237.5	34.0	1741	9 S78771	S78771 NAT=CPG 181
39	1234	33.9	37518	2 AC014003	AC014003 Drosophi
40	1234	33.9	181178	3 AC023713	AC023713 Drosophi
41	1234	33.9	306267	3 AE003442	AE003442 Drosophi
42	1211.5	33.3	14561	9 HSRING3GE	X96670 H.sapiens R
43	1211.5	33.3	38902	9 HSO27	Z96104 Human DNA s
44	1211.5	33.3	40740	9 HSO14	Z84497 Human DNA s
45	1200.5	33.0	98309	2 AL671924	AL671924 Homo sapi

RESULT 1

ALIGNMENTS

HSHUNKI
 LOCUS HSHUNKI 3149 bp mRNA linear PRI 02-MAY-1998
 DEFINITION H.sapiens HUNKI mRNA.
 ACCESSION Y12059
 VERSION Y12059.1 GI:3115203
 KEYWORDS HUNKI gene.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3149)
 AUTHORS Weber, B.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3149)
 AUTHORS Weber, B.
 JOURNAL Direct Submission
 Submitted (24-MAR-1997) B. Weber, Labor Paediatriche
 Molekularbiologie, Universitaetsklinikum Charite, Ziegelstr. 5-9,
 10098 Berlin, FRG
 FEATURES
 source location/Qualifiers
 1. 3149
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="placenta"
 223. 2391
 /gene="HUNKI"
 223. 2391
 /note="strong homology to human RING3 sequence"
 /codon_start=1
 /protein_id="CAA72780.1"
 /db_xref="GI:3115204"
 /translation="MSASGSGCTRLRLPVMGDLSTQSTTQAOAPOPANASTN
 PPEPISNPKPKQOTNOYLRLAVTLKMKHOPAPFOQPDVAVLNPDIYKIK
 TPMDGTIKRLENNYVMAOECIQDNTFTNYINKRGDDIVLAALAEKULFLK
 INELPTEETEMIVQAGRGGRKETARGVSTVNTQASTPQIQPPVPPV
 QATPPPPATVDLIVQTPVWTVPPQPLQTPPPVPPQPPAPAPVQSHPPIIA
 ATRPVTKKGVKADTPPTTIDPIHEPRLPEPKTKLGGRESSPVKPKD
 VPSOCHAPKESKSEOLKCGGILKEMPAKHAIVAMPFYVAVLALHDYCD
 IIRKPMSTIKSLARERTRDOEFGADVRLMSKCYKINPPHEVVMARKIQDYF
 EMRAKMPDEEPVAVVSSPVPPTKVAVPSSSDSSSDSSSDSSSDSSSEEA
 QRLAEOLKAVHEOLALASPOONKPKKKEKDKKKEKKEKKEVEENKSKAK
 EPPTKTKNNSSNSNKSKEPAPMKKPPPTYSEEDKCKPWSYEEKOLSIDINK
 LPEKILGVVHIIOGSEPLKNSNPDEIEIDFELKSTARELERVYTSCIRKKRKQ
 AEKVDVAGSSKMKGPFSSSESSSSSSSSSSSSSDSSSESGPA"
 BASE COUNT 818 a 946 c 788 g 597 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.92e-138 Length: 3149
 Score: 3583.00 Matches: 686
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 2
 Query Match: 98.46% Indels: 2
 DB: Gaps: 0
 US-09-700-590a-22 (1-688) x HSHUNKI (1-3149)
 QY 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 Db 223 ATGCTCGGAGAGCGCGCCCTGGAGCAGAGATTGAGAAATCTCCAGTATATGGGGATGGA 282
 QY 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 Db 283 CTAGAACTTCCCAATGCTTCAACAACAGAGCGCCCAACCCCAAGCAGCCCAACGCA 342
 QY 41 AlaSerThrAsnProProProGluThrSerAsnProAsnLysProLysArgGlnThr 60
 Db 343 GCCAGACCAACCCCCCGCCCAAGACCTTCAACCTTAACAAGCCCAAGGAGGAGACC 402
 QY 61 AsnGlnLeuGlnTyrLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAla 80
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 403 AACCACTGCATACTGCTCAGAGTGTCTCTAAGACACTATGAGAAACCACTTTGCA 462
 QY 81 TrpProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 Db 463 TGGCTTTCCAGCAGCGCTGTGATGCCGTCAAGCTGAACTCCCTGATTTACTATAGATC 522
 QY 101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGlnAsnAntyrTyrTrp 120
 Db 523 ATTAAACGCTATGATATGGGAAACATTAAGAGAGCGCTTGAGAAACCACTTATTCGG 582
 QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140
 Db 583 AATGCTCAGGAATGATATCCAGAGCTTCAACATGTTTACAAATGTTTACATCTTAAC 642
 QY 141 LysProGlyAspAspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLys 160
 Db 643 AAGCTTGAGATGATGATGATCTTAATGCGAAGAGCTTGAGAAAGCTCTTGCAAAA 702
 QY 161 IleAsnGlnLeuProThrGlnGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 Db 703 ATAAATGAGCTACCCACAGAGAAACCGAGATCATGATGCCAGCAAAAGAGAGAGA 762
 QY 181 ArgGlyLysArgLysGluThrGlyThrAlaLysProGlyValSerThrValProAsnThrThr 200
 Db 763 CGTGGAGAGAAAGAAACAGAGACAGCAAAACCTGGCGTTTCCACGGTACCAACACAACT 822
 QY 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsnProProProValGlnAla 220
 Db 823 CAAGCATGACTCTCCGCGAGACCCCAACCCCTCAGCGAATCCTCTCTGTGCAAGCC 882
 QY 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMetThr 240
 Db 883 ACGCTCAACCCCTTCCCTGCGGTCAACCCGAGACTCTCTCCAGACCCCTGTCTATACA 942
 QY 241 ValValProProGlnProLeuGlnThrProProProValProProGlnProGlnProPro 260
 Db 943 GTGATGCTCTCCCGACGACCTGCAACCCCGGCACTGCCCCCGACCAACACCCCA 1002
 QY 261 ProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThrProGln 280
 Db 1003 CCGCTCAGCTCCCGACCGCCGTCACAGCCACCCACCATCATCGGGCCACCCCAAG 1062
 QY 281 ProValLysThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle 300
 Db 1063 CCTGTGAAGACAAAGAGAGGAGTAAGAGAAAGCAACACACCCACCCACCACTATT 1122
 QY 301 AspProIleHisGluProProSerLeuProProGlnProLysThrThrLysLeuGln 320
 Db 1123 GACCCCATTCACGAGCCACCTCGCTGCCCCCGAGGCCAAGACCAACAGCTGGGCCAG 1182
 QY 321 ArgArgGluSerSerArgProValLysProProLysLysAspValProAspSerGlnGln 340
 Db 1183 CGCGGGAGAGACACCGGCTGTGAACCTCCAAAGAGAGACGCGCCGACTCCACAG 1242
 QY 341 HisProAlaProGluLysSerSerLysValSerGlnGlnLeuLysCysCysSerGlyIle 360
 Db 1243 CACCAGACACAGAGAGAGAGAGCAAGTCTCGGAGCAGCTCAAGTCTGACGCGGATC 1302
 QY 361 LeuLysGluMetPheAlaLysLysHisAlaAlaLysTrpAlaTrpProPheTyrLysProVal 380
 Db 1303 CTCAAGGAGAGTGTGCCAAGAAACACGCGCTTACCGCTGGCCCTTCTCAACACCTGTG 1362
 QY 381 AspValGluAlaLeuGlyLysHisAspTyrCysAspIleIleLysHisProMetAspMet 400
 Db 1363 GACGTGAGGACCTGGGCTTACACGACTACTGTACATCATCAAGCACCCCATGGAATG 1422
 QY 401 SerThrIleLysSerLysLeuGlnAlaArgGluLysArgAspAlaGlnGluPheGlyAla 420
 Db 1423 AGCAACATCAAGTCTTAACCTGAGGCGCGTGAACCGTATGCTCAGGAGTTTGATGCT 1482
 QY 421 AspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHisGluValVal 440
 Db 1483 GACGTCCGATTGATGTTCTCCACCTGCTATTAAGTACAACTCTCTGACATGAGAGTGTG 1542

QY 81 TrpProheGInGInProValAspAlaValLysLeuAenLeuProAspTyrTyrLysIle 100
 Db 463 TGGCTTTCAGACAGCCCTGATGCCCTCAAGCTGAACCTCCGATCTATTAAGATC 522
 QY 101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGluAsnAsnTyrTyrTrp 120
 Db 523 ATTAAGACGCTATGATATGGAACAATAAAGAGCCCTTGGAAGAACACTATTAACGG 582
 QY 121 AsnAlaGInGInLysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140
 Db 583 AATGCTCAGAGATGATATCAGAGACTTCAACATATGTTTACAAATGTTTACATCTACAC 642
 QY 141 LysProGInAspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGInLys 160
 Db 643 AAGCTGAGATGACATATGCTTAAATGACAGAGCTCGAAGAAAGCTCTTCTGCAAAA 702
 QY 161 IleAsnGluLeuProThrGluGluThrGluIleMetIleValGlnAlaLysGlyValArgGly 180
 Db 703 ATAAATAGATACCAACAGAAAGAACCGAGATCATGATATGTCAGGCAAAAGAGAGGA 762
 QY 181 ArgGlyValGlyLysGluThrGlyThrAlaLysProGInValSerThrValProAsnThrThr 200
 Db 763 CGTGGAGAGAAAGAAACAGGACAGCAAAACCTGGCGTTTCCAGCGTACCAAAACAAC 822
 QY 201 GlnAlaSerThrProProGInThrGInThrProGInProAsnProProProValGlnAla 220
 Db 823 CAAGCATGATCTCTCCGACAGACCCAGACCTCCAGCGAATCTCTCTGTGAGAGCC 882
 QY 221 ThrProHisProPheProAlaValThrProAspLeuIleValGInThrProValMetThr 240
 Db 883 ACGCTCACCCCTTCCGCGGTACCCCGGACCTCATGTCAGACCCCTGTCTCATGACA 942
 QY 241 ValValProProGInProLeuGInThrProProProValProProGInProGInProPro 260
 Db 943 GTGGGCTCTCCACAGCCACTGACAGCGCCCGCCAGTGGCCCGCCAGCCAAACCCCA 1002
 QY 261 ProAlaProAlaProGInProValGInSerHisProProIleIleAlaIleThrProGIn 280
 Db 1003 CCCGCTCAGCTCCCAACCCGCTACAGAGCCACCACTCATGTGGCGGACCCCAAG 1062
 QY 281 ProValLysThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle 300
 Db 1063 CCTGGAAGACAAAGAAAGAGTGAAGAGAAAGACACACACACCCCAACCCCAATT 1122
 QY 301 AspProIleHisGluProProSerLeuProProGluProLysThrThrLysLeuGluGln 320
 Db 1123 GACCCCATTCACAGACCACTCGCTGCCCGCCGAGCCCAAGACCAACCAAGCTGGGCGAG 1182
 QY 321 ArgArgGInSerSerArgProValLysProProLysLysAspValProAspSerGInGln 340
 Db 1183 CGCGGAGAGACAGCCGCGCTGTGAACCTCCAAAGAAAGAGTCCGACCTCTACAGAG 1242
 QY 341 HisProAlaProGluLysSerSerLysValSerGInGlnLeuLysCysCysSerGlyIle 360
 Db 1243 CACCAGACACAGAGAGAGAGGAGGCTCTGAGAGAGCTCAAGTGTGAGCGGAGATC 1302
 QY 361 LeuLysGluMetPheAlaLysLysHisAlaAlaIleTyrAlaIleProPheTyrLysProVal 380
 Db 1303 CTCAAGGAGATGTTTCCAAAGAGACAGCCGCTGAGCTTGTCTTCAAGAGCTGTG 1362
 QY 381 AspValGluAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisProMetAspMet 400
 Db 1363 GACGTGGAGGACCTGGGCTTACACAGATCTGTGACATCATTAAGACACCCCATGAGCATG 1422
 QY 401 SerThrIleLysSerLysLeuGlnAlaArgGluTyrArgAspAlaGlnGluPheGlyAla 420
 Db 1423 AGCAACAATCAAGTCTAACTGAGGCGCGTGTGATACCGTGAAGCTCAAGAGTTTGTGT 1482
 QY 421 AspValAlaGluLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHisGluValVal 440
 Db 1483 GACGCTCAGTATGATGTTCTCCAACTGATTAAGTACAACTCTGAGACATGAGGTGTG 1542

QY 441 AlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMetProAspGlu 460
 Db 1543 GCCATGGCCCGCAAGCTCCAGAGATGTTGAAATGCGCTTGGCAAGATGCCGAGAG 1602
 QY 461 ProGluGluProValValAlaValSerSerProAlaValProProProThrLysValVal 480
 Db 1603 CTTAGAGAGCCAGTGTGGCGGTCTCTCCCGCAGTGGCCCTCCACCAAGATTGTG 1662
 QY 481 AlaProProSerSerSerAspSerSerSerSerAspSerSerSerSerAspSerSerThr 500
 Db 1663 GCCCGCCCTTATCATCAGACAGACAGCAGCAGATAGCTCTCGGACATGACAGTTGACT 1722
 QY 501 AspAspSerGluGluGluValArgAlaGlnArgLeuAlaGluLeuGlnGlnLysAla 520
 Db 1723 GATGACTCTGAGAGAGAGCAGCCAGCGGCTGAGCTGAGCTCCAGAGCAGCTCAAGGCC 1782
 QY 521 ValHisGluGlnLeuAlaAlaLeuSerGlnProGlnGlnAsnLysProLysLysLysGlu 540
 Db 1783 GTGCACAGACAGCTTGAAGCCCTCTCTCAGCCCGCAGAGAAACCAAGAAAGAAAGAG 1842
 QY 541 LysAspLysLysGluLysLysLysGluLysHisLysArgLysGluGluValGluGluAsn 560
 Db 1843 AAAGACAAAGAT 1902
 QY 561 LysLysSerLysAlaLysGluProProProLysLysThrLysLysAsnAsnSerSerAsn 580
 Db 1903 AAAAAAGCAAGCCAAAGAACTCTCTTAAAAAGCAAGAAATAATAGCAGCAAC 1962
 QY 581 SerAsnValSerLysLysGluProAlaProMetLysSerLysProProProThrTyrGlu 600
 Db 1963 AGCAATGTGACAAAG 2022
 QY 601 SerGluGluGluAspLysCysLysProMetSerTyrGluGluLysArgGlnLeuSerLeu 620
 Db 2023 TCGAGAGAGAGAGCAAGTGCAGAGCTTATGTCTTATGAGAGAGAGAGAGAGAGAG 2082
 QY 621 AspIleAsnLysLeuProGluGluLysLeuGluValArgValHisIleIleGlnSerArg 640
 Db 2083 GACATCAACAAGCTCCCGGAGAGAACTGGCGCGCTGTGTGCACATCATCACTACGCG 2142
 QY 641 GluProSerLeuLysAsnSerAsnProAspGluIleGluIleAspPheGluThrLeuLys 660
 Db 2143 GAGCCCTCTCCGAAAGATTCACACCCGACAGATTAATAATCGATTGAGACCTGAAAG 2202
 QY 661 ProSerThrLeuArgGluLeuGlyAlaLeuCysHisLeuLeuPheAlaGluGluLysGlu 680
 Db 2203 CCGTCCACACTGCGTGAAGCT-GGAGCGCTATGTCACTCTCTGTGGCGAAGAAAGAA 2261
 QY 681 ThrPheLysLeuArgLysLeuMet 688
 Db 2262 ACCTC-AAAGTGAAGAAAGTTGATG 2284
 RESULT 3
 AF461396 2727 bp mRNA linear ROD 13-MAY-2002
 LOCUS
 DEFINITION Mus musculus bromodomain-containing protein BRD4 short variant
 (Brd4) mRNA, complete cds.
 ACCESSION AF461396
 VERSION AF461396.1 GI:18308126
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 2727)
 Houzelstein, D., Bullock, S.L., Lynch, D.E., Grigorieva, E.F.,
 Wilson, V.A. and Bedington, R.S.
 Growth and early postimplantation defects in mice deficient for the
 bromodomain-containing protein Brd4
 Mol. Cell. Biol. 22 (11), 3794-3802 (2002)
 JOURNAL
 MEDLINE 21993140
 PUBMED 11997514
 REFERENCE
 2 (bases 1 to 2727)

AUTHORS Houzelstein, D.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-2001) Genetique et Developpement, Institut Jacques Monod, 2 Place Jussieu, Paris 75 251, France

FEATURES
 source
 1..2727
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="17"
 /map="between D17Mit181 and D17Mit191"
 1..2727
 /gene="Brd4"
 242..2413
 /gene="Brd4"
 /codon_start=1
 /product="bromodomain-containing protein BRD4 short variant"
 /protein_id="AA167834.1"
 /db_xref="GI:18308127"
 /translation="MSTESGPGTRRLNLFVMDGLSTSQMSTTQAQAPOPANMASTN
 PPPETSNPNKPKRQTNQOYLRLVVLKTLMKHOFAMPPOQPDVAKLNLPLYKLIK
 TMDMGTILKRLNNYVMAAEOCIDQNTFMPTNCYVYKRGDDIVMAELKLFLOK
 INELPTRETEIMIYQAKGRGRKRTCTAIPGVSTVYPTQASTSPQOTPOONRPP
 VQATTHPPAVTPDLIQQPPVMTWPPQPLQTPSPVPPOPPPPAPFPQVSHPEPI
 ATTPOPVTKKVKRKAADTTPTTIDPIHEPSLAEPYAKLGPRESSREVKKPK
 DVPDSQHPGPKSKISSEQLKCCSGILKEMFAKAAVAMPFYKVDVLAJLHDYC
 DIKHPMDSTIKSLESREVDAGFADVRLPSSNDCYKYNPDHEVVAARKLODV
 FEMRFAPKMPDEBPVTVSPVAPPTVVA PPSSDSSSDSSSDSSSTDSSEER
 AORLAELQQLKAVHEQLAALSOPQONRKKKKKKKKKKKKKKKKKKKKKKKKKK
 KALPKRTKKNNSNVSKKEVPPTKTPPTPESEEDKKPKMSYSEKRLSLDTIN
 KLPGELGRVVAH IQSREPLSKNSNPDEIDIEFTLKPSTLELELYVTSCLRKRRKP
 QAEKDVIVIGSSKMKFSSSESTSESSSDSESETGPA"

BASE COUNT 802 a 763 c 636 g 526 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,02e-132 Length: 2727
 Score: 3441.50 Matches: 659
 Percent Similarity: 96.80% Conservative: 7
 Best Local Similarity: 95.78% Mismatches: 21
 Query Match: 94.57% Indels: 3
 DB: 10 Gaps: 1

US-09-700-590A-22 (1-688) x AF461396 (1-2727)

QY 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 DB 242 ATGTCTACGAGAGCGCCCTGGGACAAAGATTGACAGTAATGGGAGATGGA 301
 QY 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 DB 302 CTGAAACCTCCCAATGTGTACACCGAGGCCCAAGCCCAACCCAGCAGCAATGCA 351
 QY 41 AlaSerThrAsnProProProGlyThrThrSerAsnProAsnLysProLysArgGlnThr 60
 DB 362 GCCAGACCAATCTCTCCACCCCGAGAGACTCCAACTTAACAAGCCCAAGAGACAGACA 421
 QY 61 AsnGluLeuGlnThrLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnIleAla 80
 DB 422 AACCAACTGCAATATCTGCTCAGAGTGGTCTCAAGCACTATGAGAAACCCAGTTTGGC 481
 QY 81 TrpProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 DB 482 TGGCTTTTCCAGAGAGCCCGTGATGCCGTCAAGCTCACTCTGATTTACTTAAGATT 541
 QY 101 IleLysThrProMetAspMetGlyThrIleLysAlaArgLeuGlnAsnAsnTyrTyrTrp 120
 DB 542 ATTAATAACACCCATGATATGGAAACAATAAGAGCGCTTGAATAAACAATACTATGCG 601
 QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnGlyTyrIleTyrAsn 140
 DB 602 AATGCTCAGGAATGTATCCAGACTTCACACTATATGTTTAAATAATGTATCACTATTAAC 661

QY 141 LysProGlyAspAspIleValLeuMetAlaGluAlaLeuGlnLysLeuPheLeuGlnLys 160
 DB 662 AACCTCGAGATGATCATCTCTTAATGGAGAGAGCTCTGGAGAAAGCTCTTCTTGCAAAAA 721
 QY 161 IleAsnGluLeuProThrGlnGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 DB 722 ATCAATGAACCTGCTACAGAAAGAACTGAGATCATATGATCTCAGGCAAGAGAAAGAGA 781
 QY 181 ArgGlyArgLysGlnThrGlyThrAlaLysProGlyValSerThrValProAsnThrThr 200
 DB 782 CGAGGAGAGAAAGAAACAGAGACAGCAAGAGCTGTGTATCCACGGTACCACCAACAAC 841
 QY 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsn---ProProProValGln 219
 DB 842 CAAGCATCAACTCTCTCCGAGAGCCAGAGCGCTCAGCAACAACCTCTCCACTGTGCAG 901
 QY 220 AlaThrProHisProPheProAlaValThrProAsnLeuIleValGlnThrProValMet 239
 DB 902 GCCACAACCTCACCCCTTCTCTGTCTCACCCAGACCTATGCTCCAGCTCTGTCAATG 961
 QY 240 ThrValAlaProProGlnProLeuGlnThrProProProValProProGlnProGlnPro 259
 DB 962 ACAATGTGCCCCCTTACAGCCACTTTCAGACTCTTCAACGGTACCCCCCAGCCACACACC 1021
 QY 260 ProProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThrPro 279
 DB 1022 CCACCTGTCTCAGATTCCACAGCCTGTGCAGAGTCAACCCGCCCATCTTGTGCACACCCC 1081
 QY 280 GlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThr 299
 DB 1082 CAGCCCTGAAGCAAAAGAAAGGGGGAAGAGAAAGCAAGATACCAACCCCTTACACCC 1141
 QY 300 IleAspProIleHisGluProProSerLeuLeuProProGlnProLysThrThrLysLeuGly 319
 DB 1142 ATGACACCCATTCATAGGACCACTCTACCTGAGCCCAAGACCCCAAGCACTGGGT 1201
 QY 320 GlnArgArgGluSerSerArgProValLysProProLysLysAspValProAspSerGln 339
 DB 1202 CCTCGGCGGAGAGAGCAGACCTGTGAAGCCTCCAAAGAGAGTGTACGGACTCACAG 1261
 QY 340 GlnHisProAlaProGlnLysSerSerLysValSerGlnGlnLeuLysCysCysSerGly 359
 DB 1262 CAGCACCCAGGGGCAGAGAGAGAGCAAGATCTCGACACCTAAAGTGTGCACTGAC 1321
 QY 360 IleLeuLysGluMetPheAlaLysLysHisAlaAlaTyrAlaTyrProPheTyrLysPro 379
 DB 1322 ATCTTAAGAGAGATGTTTCTCCAAAGAAACATGCTGCTATGCTTGCCCTTTCTACAGCCT 1381
 QY 380 ValAspValGluAlaLeuGlnLysLeuHisAspTyrCysAspIleIleLysHisProMetAsp 399
 DB 1382 GTGGATGTGAGAGCACTGGGTCTGCACGACTACTGTGACATCATCAACATCCCATGGAGC 1441
 QY 400 MetSerThrIleLysSerLysLeuGlnAlaArgGluTyrArgAspAlaGlnGluPheGly 419
 DB 1442 ATGAGCACAATCAAGCTCAACTAGAGTCCGAGAGTCAAGAGATGCCAGAGAAATTTGAT 1501
 QY 420 AlaAspValArgLeuMetPheSerAsnGlyTyrLysTyrAsnProProAspHisGluVal 439
 DB 1502 GCTGATGTCGATTGATGTTCTTCCACTGCTACAGTACAAACCCCTTACATGAAGTG 1561
 QY 440 ValAlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMetProAsp 459
 DB 1562 GTAGCCATGGCTGAAATACTCAAGATGTGTTGAATGCGTTTGCCAAAGTGGCTGAT 1621
 QY 460 GluProGluGluProValValAlaValSerSerProAlaValProProProThrLysVal 479
 DB 1622 GAGCTGGAAGAGCCAGTTGTTACAGTGTCTCTCTGCAAGTCCACCCCTTCAAAAGG 1681
 QY 480 ValAlaProProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 499
 DB 1682 GTAGCCCAACCTCATCTAGTACAGACAGCAAGCAAGTGTCTTCCAGACGCAAGTTC 1741
 QY 500 ThrAspAspSerGluGlnGluArgAlaGlnArgLeuAlaGluLeuGlnGlnLeuLys 519

D	b	601	AAGCTGGAATGCATGTCTTAAATGGCAGAAGCTTGGAAGAAAGCTTTCTTGCAAAA	660
O	y	161	IIeaNGIueuPProThrGIuGIuThrGIuLIeMeIIeValGIuaLySGIyArGIy	180
D	b	661	ATCAATGAATGCCCTACAGAAAGAAATCGAATCATGTATGTCACAGGCAAAAGAAAGAGA	720
O	y	181	ArgGIyArGIySGIuThrGIyThrAlaLySProGIyValSerThrValProAntThr	200
D	b	721	CGAGGAGAGAAAGAAACGAGGGGACAAAGCGTGTAATCCAGGTAAACAACAACCT	780
O	y	201	GInaLSeIThrPrProGIuThrGIuThrProGIuProSen---ProProProValGIu	219
D	b	781	CAGCATCAACTTCTCCGCAAGACCAGCGCTCAGAGAAACCTCTCACCTGTGCGAG	840
O	y	220	AlaThrProHISProPheProAlaValThrProAspLeuIIeValGIuThrProValMe	239
D	b	841	GCCAAACTCAACCCCTTCTGTGTGTCAACCCAGACCTCATATTGCCAGCCTCTGTCAATG	900
O	y	240	ThrValValProProGIuThrProLeuGIuThrProProProValProProGIuThrPro	259
D	b	901	ACAATGGGCCCCCTCAGCCACTTCAGACTCTTCACCGGTACCCCCCGACCAACACC	960
O	y	260	ProProAlaProAlaProGIuThrProValGIuSerHisProProIleIIeAlaIaThrPro	279
D	b	961	CACTGTCTCAAGTTCCACAGCCTGTGCAGAGTCACCCGCCATCATATGGACACCCC	1020
O	y	280	GInProValLySThrLySGLyValLySArGIySaLySAlaSPThrThrThrProThrThr	299
D	b	1021	CAGCCTGTGAAGACAAAGAAAGGGGTGAAGAGAAACAATACACACCCCTTACACAC	1080
O	y	300	IleaSProlIeHISGIuProProSerLeuProProGIuProLySThrThrLySLeuGIy	319
D	b	1081	ATCGACCCCATTCATGTAGCCACCTCACTAGGCCCAAGGCCAAAGAACCGCAAGCTGG	1140
O	y	320	GInaPrgArGIySerSerArGIyProValLySProProLySAspValProAspSerGIu	339
D	b	1141	CCTGGCGGGAGACAGACAGCTGTGAAGCCTCAAGAAAGAGATGACCGGACTACAG	1200
O	y	340	GInHISProAlaProGIuLySserSerLySValSerGIuInLeuLySCysCysSerGIy	359
D	b	1201	CAGACCCAGGGCCAGAGAGAGACAGACAATCTCTGAGCAGCTAAAGTCTGCAATGCC	1260
O	y	360	IleuLySGIuMetPheaLySLeuYHisAlaIaLyTaIaTrpProPheTyLySPro	379
D	b	1261	ATCCTCAAGGAGATGTTGGCCAAAGAAACATGTGCTATGCTGGCCTTTCTTCAAGCCT	1320
O	y	380	ValSerValGIuAlaLeuGIyLeuHISarPTyrCyAspRIleIleLySHISProMetLar	399
D	b	1321	GTGATGTGGAGGACGTGGCTGTGCAGAGTACTGTGACATCATCAACAACCTCCAGAGC	1380
O	y	400	MetSerThrIleLySserLySLeuGIuAlaArgGIuTyrTrpAspAlaGIuInPuPheGIy	419
D	b	1381	ATGAGCAATCAATCAAGCTTAAACTTAGAATCCCGAGAGTACAGAAATGCCAGAAATTGG	1440
O	y	420	AlaAspValaArgLeuMetPheSerAsnCyTyLySTyrAnbProProAspHISGIuVal	439
D	b	1441	GCTGATGTCCGATGATGTCTTCCAATGTGTACAGTACAACCCCTGATCCATGAAGTG	1500
O	y	440	ValAlaMetAlaArgLySLeuGIuInasPvalPheGIuMetLargPheAlalyShetProAsp	459
D	b	1501	GTACCCATGGCTCCAAAACTCCAGAGATGTTTGAATGGCTTTCCCAAAGTCCGTGAT	1560
O	y	460	GIuProGIuGIuProValAlaIaValaSerSerProAlaValProProProThrLySVaI	479
D	b	1561	GAGCTGAAGGCAAGTTTTCACAGTGTCTCTCTCTGCAATGCCACCCCTTCAAAAGTG	1620
O	y	480	ValAlaProProSerSerSerAspSerSerSerSerAspSerSerSerSerAspSerSer	499
D	b	1621	GTAGCCCAACCTCATCATAGTGAACAGACAGCAAGTTCTTCCAGACAGCAAGTTCC	1680
O	y	500	ThrAspAspSerGIuGIuGIuArgAlaGIuArgLeuAlaGIuLeuGIuGIuGIuLeuLyS	519
D	b	1681	ACTGACACTCTGAGGAAGGACAGCCAGCGGCTGTGAATCTCAGAAACAGCTCAAG	1740
O	y	520	AlaValHISGIuInLeuAlaAlaLeuSerGIuProGIuInGIuAsnLySProLySlyS	539
D	b	1741	GCCGTGATGAGCAGCTTGTGAGCCCTCTCTACAGCCCCCAGCAACAAACCAAAGAAAA	1800
O	y	540	GIuLySAspLySGLyGIuLySlySGLyGIuLySHISLySArgLySGIuValGIuGIu	559
D	b	1801	GAGAAAGACAAAGAAAGAAAGAAAGAAAGAACACAAAGAAAGAAAGATGGAGAA	1860
O	y	560	AsnLySlySserLySAlaLySGIuProProPolylLySThrLyLSAsnAsnSerSer	579
D	b	1861	AATATAAAAGCAAAACCAAGAACTCTCTCCAAAAAGACAAAGAAAAATACAGCAGC	1920
O	y	580	AsnSerAsnValSerLySlySGIuProAlaPrometLySserLySProProThrTyr	599
D	b	1921	AACAGCAATGTGACAAAGAAAGAACAGTACCCAGAAAGAACCGCCTCCACATAT	1980
O	y	600	GIuSerGIuGIuGIuAspLyScyLySPrometSerTYrGIuGIuLySArgGIuLeuSer	619
D	b	1981	GAATCAGAAAGAGAGTAAGTGAAGCCCATGTCTTATGAGAGAAAGCGCAGCTAAGT	2040
O	y	620	LeuAspIIeAsnLySLeuProGIyGIuLySLeuGIyArgValValHISIIeIIeGIuSer	639
D	b	2041	CTGATATTCACAAACTCTCTGTGTGAAGATTAGCCCGTGTGTACATTAATTAGTCA	2100
O	y	640	ArgGIuProSerLeuLySAsnSerAsnProAspGIuIIeGIuIIeAspPheGIuThre	659
D	b	2101	AGGGAACCATCACTTAAAACTCCAAACCCGATGAGATTGAGATTGACTTGTGACCCCTG	2160
O	y	660	LysProSerThrLeuArgGIuLeuGIyAlaLeuCyGHISLeuLeuPheaAlaGIuGIuLyS	679
D	b	2161	AAGCCATCTCACTTACAGAGATT-GGAGGGAATATGTCACTCTGTTTGGAGAAAGAA	2219
O	y	680	GIuThrPheLySLeuArgLySLeu 687	
D	b	2220	GAATCTTC-AAAGCTGAAGAAAGTTG 2242	
RESULT 5				
AF461395 LOCUS 6000 bp mRNA linear ROD 13-MAY-2002				
DEFINITION Mus musculus bromodomain-containing protein BRD4 long variant				
(Brd4) mRNA, complete cds.				
ACCESSION AF461395				
VERSION AF461395.1 GI:18308124				
KEYWORDS Mus musculus.				
SOURCE Mus musculus.				
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE 1 (bases 1 to 6000)				
AUTHORS Houzelstein,D., Bullock,S.L., Lynch,D.E., Grigorjeva,E.F., Wilson,V.A. and Beddington,R.S.				
TITLE Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4				
JOURNAL Mol. Cell. Biol. 22 (11), 3794-3802 (2002)				
MEDLINE 21993140				
PUBMED 11997514				
REFERENCES 2 (bases 1 to 6000)				
AUTHORS Houzelstein,D.				

```

/codon_start=1
/product="Bromodomain-containing protein BRD4 long
variant"
/db_xref="GI:18308125"
/translation="WSTESGEGTRRLNLPVWGGLGTSQNSTTQAOQOPAPANAStN
PPEPESNPKNPKRQTNQQLRLVLTLMKQFAPFOQPDVAVLNPDYKILK
TPMDMGTIKRLENNYWNAOECIODFTNCTYINKGQDVLMAEALKEFLQK
INELPTEETIMIVQAKRGGRGKGTGAARGVSTVNTQASTSPOTQOPNPPI
VQATTPPEPAVTPDLIAOPVMTVPPOPTVPPVPOPPVPOPPVPOPPVPOPPV
ATTPOPKTKKQKAKADTTTPTTPIPIHPSPASPPAKIKAPRESRRPPKPK
DVPDSQOHPPEKSSKTSBQKCCSGILKMFKAHAIVAPPTKPDVAGLHDYC
DI1KHPMDSTIKSKLESREYDAOEQAVRMLFNSCYXNPPRDHVVAMARKLOV
FEMRFAPDEPEEPVVTSSAPVPTKASVSSDSDSDSDSDSDSDSDSDSDSEER
AOLRLAEQOLKAVHQAALSOQONPKKEKEDKKKKKKKKKEEVENKSKTK
KELPPKTKNNSSNSVSKKEPVPKTPPTYESDEEKKCKMSEERKQSLDIN
KLPEKLGAVVHTIIOREPLKSNPDEIETDEFTLRLELRVYVTSCLKRRR
QAEKVDVIASSSKMKFSSSESTSSSDSDSDSTEMAPSKKKGHTGRDQKH
HHHPQOPAPAPVPOPPPPPOPPPOPPPOPPPOPPPOPPPOPPPOPPPOPP
PPEPTAQVLEPOLGVSVDPISTFOTPIHLPOBELPHILPOPEHSTPHLNH
AVSPRALHNALPOQSPRSNRALAPKPTRPAPVAPALAPOLLPQPMAPPOVL
LEDEPAPPLUTSMQOLYLOQLOKQVOPPTPLPSVVOQOPPPPLPPPHSVQOO
LOQPPPPPPPOPPPOQOHPPPRVNHPMSRPAHIOQPPPPPOQPPPPPOPP
PPOPPAPQOVIOHPSPRHNSDPTSAQHLRABPLWTHSQMFOSQSLTQSP
QONVQPKQVKGRAEPQPGVWGQGGCPASPAAPMLSOELRPSVVOQPLVV
KEEKHSPIIRSEPFSTSLRPPKHEENIKAPHLPOPEMKVDIGRIVIRPEOS
APPKADKDKQKQEPKTPVAPKDKLIKMSWSASVQKHPTTPTSSSTAKSSPFEH
FRARAREKERELKAOAHAEKEKRLQERHRSDEDEALQARABEARROE
QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO
MAATIDNPFQSDLLSIFENLF"
BASE COUNT      1624 a      1835 c      1380 g      1161 t
ORIGIN
Alignment Scores:
Pred. No.:      3,336-132      Length:      6000
Score:          3436.50      Matches:      658
Percent Similarity: 96.66%      Conservative: 7
Best Local Similarity: 95.64%      Mismatches: 22
Query Match:    94.44%      Indels:      3
DB:             10      Gaps:      1
US-09-700-590A-22 (1-688) x AF461395 (1-6000)
QY      1 MetSerIaGluSerGlyProGlyThraGleuAArgAneLeuProValMetGlyAspGly 20
Db      242 ATGCTTACGAGAGAGCGCCCTTGGAACAAGATTGAGAATCTGCCAGTAATGGGGATGA 301
QY      21 LeuGluThrSerGluMetSerThrThrglnaGlnaGlnaGlnProGlnProAlaAsnAla 40
Db      302 CTAGAAACCTCCCAATATGTCACAGCAGCGCCCAACCCCAAGCCAGCAAGCAATGCA 361
QY      41 AlaSerThraAnProProProGluThrSerAnProAsnLeuProLysArgGlnThr 60
Db      362 GCCACACCAATCTCCCAACCCCAAGAGACTCAACCTCAACAAACCCCAAGAGACAGCA 421
QY      61 AsnGlnLeuGlnThyLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAla 80
Db      422 AACCAACTGCATATCTGCTCAGAGTGTGCTCAAGACACTATGAAACACCAAGTTTGG 481
QY      81 TrpProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
Db      482 TGGCCTTCCAGCAGCCCGTGATGCCGTCAGCTGAACCTCCCTGATTAATAAGATT 541
QY      101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGlnAspAsnTyrTyrTrp 120
Db      542 ATTAAACACCAATGATGATATGGAAACAATTAAGAGCGCTTGAAACCACTATTAATCTGG 601
QY      121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140
Db      602 AATGCTCAGGAATTAATCCAGACTTCAACACTATGTTTACAAATTTGTTACATCTATAAC 661
QY      141 LysProGlyAspAspIleValLeuMetAlaGlnaIleGluLysLeuPheLeuGlnLys 160

```

```

Db      662 AAGCTCGAGATGACATCGCTTAATGGCAAGACTCTGAGAAAGCTCTTTCGCAAAA 721
QY      161 IleAsnGluLeuProThraGluGluThraGluIleMetIleValGlnAlaLysGlyArgGly 180
Db      722 ATCATATGAATCGCTTACAGAAAGAACTGAGATCATGATATGTCAGGCAAAAGAAAGAGA 781
QY      181 ArgGlyValGlyGluThraGlyThrAlaLysProGlyValSerThrValProAsnThrThr 200
Db      782 CGAGGGGAGAAAGAAACAGGGGAGCAAAACCTGCTGTATACCAAGTACCAAAACAACT 841
QY      201 GlnAlaSerThrProProGluThraGlnThrProGlnProAsn---ProProProValGln 219
Db      842 CAAGCATCAACTCTTCGCAAGACCCAGACCTTCAGAGAAACCTCTCAACTGTGACAG 901
QY      220 AlaThrProHisProPheProAlaValAlaThrProAspLeuIleValGlnThrProValMet 239
Db      902 GCCCAACTACACCTTCTCTGCTGTACCCAGACCTCATTTGGCCAGCGCTCGTATG 961
QY      240 ThrValValProProGlnProLeuGlnThrProProProValProProGlnProGlnPro 259
Db      962 ACAATGGTGGCCCTCAGCCACTTCAGACTTCACCTTCACCCGTTACCCCCCAAGCAACACC 1021
QY      260 ProProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThrPro 279
Db      1022 CCACCTGCTCAGTTCCACAGCTGTGCAAGATCACCCGCCCATATTGGACCAACCCCC 1081
QY      280 GlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThr 299
Db      1082 CAGCTGTGAAGAAAGAAAGAAAGGGGTGAAGAGGAACAGATACCAACACCCCTACACC 1141
QY      300 IleAspProIleHisGluProProProSerLeuProProGluProLysThrThrLysLeuGly 319
Db      1142 ATCAACCCCAATTCATGAGCCACCTCCTGACCCCAAGCCCAAGACCCCAACCTGAGT 1201
QY      320 GlnArgArgGluSerSerArgProValLysProProLysLysAspValProAspSerGln 339
Db      1202 CCTGGGGGGAGACAGACAGACTGTGAAGCTTCAAAGAAAGATGACCGGACTCACAG 1261
QY      340 GlnHisProAlaProGluLysSerSerLysValSerGlnLeuLysCysCysSerGly 359
Db      1262 CAGCACCCAGGGCCAGAGAAAGAGCAGCAGCATCTGAGCAGCTAAAGTGTGTCAGTGGC 1321
QY      360 IleLeuLysGluMetPheAlaLysLysHisAlaAlaTyrAlaTrpProPheTyrLysPro 379
Db      1322 ATCTCAAGAGAGATGTTGGCCAAAGAAACATGCTGCTGCTGCTGCTTCTCAACACCT 1381
QY      380 ValAspValGlnAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisProMetAsp 399
Db      1382 GTGATGTGGGSCACTGGGTCTGCAAGACTACTGTACATCATCAAAACATCCCATGGAC 1441
QY      400 MetSerThrIleLysSerLysLeuGlnAlaArgGluLysTrpArgAspAlaGlnGluPheGly 419
Db      1442 ATGAGCAATCAAGTCTTAACCTAAGTCCCGAGAGTCCGAGAGATGGCCAGGAATTTGGT 1501
QY      420 AlaAspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHisGlnVal 439
Db      1502 GCTGATGCTCCATGATGTTCTCCAACTGCTACAAAGTAAACCCCTGACCAATGAAGTG 1561
QY      440 ValAlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMetProAsp 459
Db      1562 GTAGCCATAGGCTCAAAAACCTCAGAGATGTTTGAATAGCGCTTTGGCAAGATCCCTGAT 1621
QY      460 GluProGluGluProValAlaAlaValSerSerProAlaValProProProThrLysVal 479
Db      1622 GAGCTGAAGAGCCAGTTTGTACAGTGTCTCTCTGCAAGTGCACCCCTACAAAGGTG 1681
QY      480 ValAlaProProSerSerSerAspSerSerSerSerSerSerSerSerSerSerSerSer 499
Db      1682 GTAGCCCAACCTCATCTAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1741
QY      500 ThrAspAspSerGluGluGluArgAlaGlnArgLeuAlaGluLeuGlnGlnGlnLys 519
Db      1742 ACTGACGACTCTGAGAGAAAGACCGACCGCGCTGACTGAATCTCAGGAACACCTCAAG 1801

```

QY 520 AlValHISGLIUGLInLeuAAlaLeuSerGlnProGlnGlnAnLysProLysLysLys 539
 DB 1802 GCCGTGATGAGCAGCTTGACGCCCTCTCCAGCCCGAGCAAGCAAAACCAAGAAAAAG 1861
 QY 540 GILUYSAPLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 559
 DB 1862 GAGAAAGCAAGAAAGAAAGAAAGAAAGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1921
 QY 560 AsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 579
 DB 1922 AATTAATAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1991
 QY 580 AsnSerAnValSerLysLysLysLysLysLysLysLysLysLysLysLysLysLys 599
 DB 1982 AACAGCAATGTGAGCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2041
 QY 600 GILUYSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 619
 DB 2042 GAATCAGAAAG 2101
 QY 620 LeuAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 639
 DB 2102 CTGATATTCACAAACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2161
 QY 640 ArgGluProSerLeuLysAsnSerAnProAspLysLysLysLysLysLysLysLys 659
 DB 2162 AGGAAACATCATTAATAAACTCCACCCGATGAGATTGAGATTGAGATTGAGATTGAG 2221
 QY 660 LysProSerThrLeuAAGLysLysLysLysLysLysLysLysLysLysLysLysLys 679
 DB 2222 AAGCATCTACACTACAGAGATT-GGAGCGATATGATGATGATGATGATGATGATGAT 2280
 QY 680 GluThrPheLysLeuArgLysLeu 687
 DB 2281 GAAACCTC-AAGCTGAAAAAGTTG 2303

RESULT 6
 LOCUS BC008532 2199 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, clone IMAGE:3489640, mRNA.
 ACCESSION BC008532
 VERSION BC008532.1 GI:14250219
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2199)
 STRAUSBERG, R.
 Direct Submission
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Lohar Hemmighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-ehg.stanford.edu>
 Contact: (Dickson, Mark) med@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

REMARK
 COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 7 Row: b Column: 12.
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF

FEATURES
 source analysis, Similarity but not identity to protein.
 Location/Qualifiers
 1..2199
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C57BL/6J"
 /clone="IMAGE:3489640"
 /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."
 /clone_1lb="NCI CGAP_Mams"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 BASE COUNT 611 a 642 c 542 g 404 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,22e-96 length: 2199
 Score: 2556.50 Matches: 481
 Percent Similarity: 96.63% Conservative: 6
 Best Local Similarity: 95.44% Mismatches: 16
 Query Match: 70.25% Indels: 1
 DB: 10 Gaps: 1
 US-09-700-590A-22 (1-688) x BC008532 (1-2199)
 QY 48 ProGluThSerAnProAnLysProLysArgGlnThrAsnGlnLeuGlnLysLeu 67
 DB 685 CCAGAGACCTCCCAACCTTAACAAGCCCAAGAGACAGACAAACCAATGCAATCTGCTC 744
 QY 68 ArgValAlaLeuLysThrLeuTpyLysLysLysLysLysLysLysLysLysLysLys 87
 DB 745 AGAGTGGTGTCAAGACCTATGGAACACACACTTGGCTGGCTTCCAGAGCCCGG 804
 QY 88 AspAlaValLysLeuAnLysLeuProAspTyrTyrLysLysLysLysLysLysLysLys 107
 DB 805 GATGCCGTCAAGCTGAACCTCCCTGATTAATTAATTAATTAATTAATTAATTAATTA 864
 QY 108 GILYThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 127
 DB 865 GGAACATTAAGAGAGCGCTTGAGAAACAATTAATTAATTAATTAATTAATTAATTA 924
 QY 128 AspPheAnThrMetPheThrAnCysTyrLysLysLysLysLysLysLysLysLysLys 147
 DB 925 GACTTAACACATTAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 984
 QY 148 LeuMetAlaGlnAlaLeuGlnLysLysLysLysLysLysLysLysLysLysLysLys 167
 DB 985 TTATGGCAGAGCTTGAGAGAGCTTCTTGCAAAAATCAATGATGCTCTACAGAA 1044
 QY 168 GILUThGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 187
 DB 1045 GAAACCTGAGATCATGATGATCCAGGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1104
 QY 188 ThrAlaLysProGlyValSerThrValProAsnThrThrGlnAlaSerThrProGln 207
 DB 1105 GCAGCAAAAGCTGTATGATCCAGGTAACAACAACAACAACAACAACAACAACAACA 1164
 QY 208 ThrGlnThrProGlnProAnLysLysLysLysLysLysLysLysLysLysLysLys 226
 DB 1165 ACCCAAGCCCTCAGAGAACCTCTCCACCTGTCAGAGCAACAACAACAACAACAACA 1224
 QY 227 AlAlaValThrProAspLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 246
 DB 1225 GCGTGAACCCAGACCTTATGCGCCAGCTCTGTCATACATAGATGATGATGATGAT 1284
 QY 247 LeuGlnThrProProProValProProGlnProProGlnProProProValProProGln 266
 DB 1285 CTTCAGACTCTTCAACCGGTACCCCGCCAGCACCAACCCCACTGCTCAGATTCCAG 1344
 QY 267 ProValGlnSerHisProProLysLysLysLysLysLysLysLysLysLysLysLysLys 286
 DB 1345 CCGTGCAGAGTCAACCCGCAATTCATTCGACCAACCCCGCACTGTGAGAAAGAA 1404

```

QY 287 GlyValLysArgLysAlaAspThrThrThrProThrThrIleAspProIleHisGluPro 306
Db 1405 GGGGTGAAGGAAGCAAGATACCAACCCCTACCAACCATCGACCCCATTCATGAGGCA 1464
QY 307 ProSerIeuProProGluProLysThrThrLysLysGluIleAspArgGluSerSerArg 326
Db 1465 CCCCTACGCGCCCAAGACCCCAAGACCCCAAGCTGGGTCTCCGGGGAAGACACGCA 1524
QY 327 ProValLysProProLysLysAspValProAspSerGlnGlnIleAspProGluLys 346
Db 1525 CCTTGAAGCTCCCAAGAAAGATGACCGACCTCACAGCAGCAGCCGAGGCGCAGAGAG 1584
QY 347 SerSerLysValSerGluGlnLysCysSerGlyIleLysGluLysGluMetPheIa 366
Db 1585 AGCAGCAAGATCTCGAGCAGCTAAAGTGTGAGTGGCATCCCAAGAGATGTTGCC 1644
QY 367 LysLysHisAlaAlaIleAlaIleProPheThrLysProValAspValGluAlaLysGlu 386
Db 1645 AAGAAACATGCTGCTTGAAGCTTGGCTTTCACAAAGCTGATGTTGAGGCACTGGGT 1704
QY 387 LeuHisAspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLys 406
Db 1705 CTGCACGACCTACTGTGACATCATCAACATCCCATGAGCATGAGCAGCATCAAGTCTAA 1764
QY 407 LeuGluAlaArgLysArgAspAlaGlnGluPheGluAlaAspValArgLysMetPhe 426
Db 1765 CTAGAGTCCCGAGAGTACAGAGATGCCAGGAATTTGGTGTGATGTCGATGATGTTTC 1824
QY 427 SerSerCysTyrLysTyrAspProProAspHisGluValIleAlaMetAlaArgLysLys 446
Db 1825 TTCACATGCTCAAGATACACCCCTGACCAAGAGTGAGTACCATGCTGCAAACTC 1884
QY 447 GlnAspValPheGluMetArgPheAlaLysMetProAspGluProGluLysProVal 466
Db 1885 CAGAGTGTGTTGAATTCGGCTTTCACAAAGTCTGATGAGCCTGGAAGAGCAGTGTGT 1944
QY 467 AlaValSerSerProAlaValProProThrLysValIleAlaProProSerSerSer 486
Db 1945 ACAGTGTCTCTCCGACGAGTCCACCCCTACAAAGGTGAGGCCCACTCATCTAGT 2004
QY 487 AspSerSerSerAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 506
Db 2005 GACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2064
QY 507 ArgAlaGlnArgLysAlaGluLysGlnGlnLysLysAlaValHisGluGlnLysAla 526
Db 2065 CGAGCCGACGCGCTGCTGACATCCAGGAACGCTCAAGCGCGCTGACAGAGAGCTTCA 2124
QY 527 AlaLysSerGlnProGlnGlnLysProLysLysGluLysAspLysLysGluLys 546
Db 2125 GCGCTCTCAAGCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2184
QY 547 LysLysGluLys 550
Db 2185 AAAAAAAAAA 2196

```

RESULT 7
BC031536
LOCUS BC031536 3033 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, bromodomain-containing 3, clone MGC:28328
IMAGE:4015879, mRNA, complete cds.
ACCESSION BC031536
VERSION BC031536.1 GI:21594630
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3033)
REFERENCE Direct Submission
TITLES Submitted (06-JUN-2002) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer

```

REMARK
COMMENT
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gamaratne, P.H., Garcia, A.M., Lu, X., Huliyil, S.W., Hale, S.M.,
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LUNL at: http://image.lnl.gov
Series: IRAP Plate: 36 Row: m Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12363646.
1..3033
location/Qualifiers
source
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZCH 11"
/clone="MGC:28328 IMAGE:4015879"
/tissue_type="Mammary tumor metastasized to lung.
MMTV-LTR/Mnt1 model. Expression driven by an MMTV-LTR
enhancer."
/clone_1ib="NCI CGAP Lu30"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
167..2347
/codon_start=1
/product="bromodomain-containing 3"
/protein_id="AAH31536.1"
/db_xref="GI:21594631"
/translacion="MSTTAAATGTPAVGPNVPPEPVSNSPKGRKTNOLOMONV
VVKTLKHOPAMPFYQVDAIKLNLIDYKILINPMDSGITIKRLNNYWSASBCMO
DEPTMTNCTYINKEPTDVLVMAOLEKIFLOKVAOMPOBEVILPAPGSKRPA
GAONASQOVAIVSSVSPATPNTPTPTYSOTVIAATVPTTAVTSVPPPPAP
PPATPITVPPVPPVKKGVKARADITPTTSAITLSSSSPPLSEPKAKVVA
RRSGGRPIKPKKDLDEVDPOHAKKGLSLHRCOSILLEMSSKHAAYAMPY
KPYDABARLELDHDIKIPMDISTYRKMDSEHYDPAOGFADIRLMSNCYKNP
DHEVVMARLQDFEMRPAKMPDEMEAPALPAPAPIVSKGABSSRSSESSDSG
SDSEBERARLAEILOELKAYHEOLAAISOAVNPKKKKKKKKKKKKKKKKK
KEHGKAKSEBEKAKAPPAKAOQOKKAPTKKANSNTTASROLKGGKQASASYS
EEGGLPWSYDEKQSLDINRLPDEKLGKRVYHIOGRBSIRDSNPDEIFIDEETIKP
TTLRELERYVKSCLQKQKRPSTSGKKQAKSKELADQKKELEKRLQDVSGQLNS
KKPTKKKSGSAPSGSPSRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS"
BASE COUNT 844 a 827 c 832 g 530 t
ORIGIN
Alignment Scores:
Pred. No.: 5,39e-71 Length: 3033
Score: 1931.00 Matches: 412
Percent Similarity: 69.05% Conservat: 70
Best Local Similarity: 59.03% Mismatches: 154
Query Match: 53.06% Indels: 63
DB: Gaps: 18
US-09-700-590A-22 (1-688) x BC031536 (1-3033)
QY 9 ThrArgLeuAlaGlnLeuProValMetGlyAspGly-----LeuGluThrSerGlnMet 26
Db 110 ACAGAGACCGCCTCTTGAGCTGCTGTGACAGCAAGTCACTAAGCTGCCAGAGATG 169
QY 27 SerThrThrglnAlaGlnAlaGlnProGlnProAlaAsnAlaIleSerThrAsnProPro 46
Db 170 TTCACCTACAGCGGCTGCCCAAGGAGATCCCGGACATGCCCGGAGCCCGCTGACCTCC 229

```

[illegible]


```

Db 1341 AGAAGCTCCAGATGTTGAGATGAGCTTTGCCAAGATGCCCATGAGCCCATGAG 1400
Qy 464 ProValValAlaValSerSerProAlaValProProThryValValAlaProPro 483
Db 1401 -----GGGCTGGGCTGCCGCTCCAGGCGCCCATCGAGAGCAAG 1442
Qy 484 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerThraPaspSer 503
Db 1443 GGGGCTGAGAGCAGCCGAGTAGTAGAGAGAGCTCTTCAATTACAGGACCTGACTCA 1502
Qy 504 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 523
Db 1503 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
Qy 524 GlnLeuAlaAlaLeuSerGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 542
Db 1563 CAGCTGGATGCTTGTCTCTCAGGCCCCAGTGAACAACCAAGAGAGAGAGAGAGAG 1622
Qy 543 LysLysGluLysLys-----LysGluLysLysLysLysLysLysLysLysLys 555
Db 1623 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAGGCC 1682
Qy 556 GluValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 571
Db 1683 AAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
Qy 572 Lys-----ThryLysLysLysLysLysLysLysLysLysLysLysLysLysLys 589
Db 1743 AAGGCTCCCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1796
Qy 590 PrometLysSerLysProProProThryThryGluSerGluGluGluGluGluGlu 609
Db 1797 AAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1856
Qy 610 MetSerThryGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 629
Db 1857 ATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1916
Qy 630 LeuGluArgValValHisIleIleGlnSerArgGluProSerLeuLysSerSerPro 649
Db 1917 CTGAGGCGGTGTGTGCATCATTCAGTCTCGGAGAGCCCTCGCTTCGGAGCTAAACCA 1976
Qy 650 AspGluIleGluIleAspPheGluThrLeuLysProSerThryLeuArgGluLeuGlyAla 669
Db 1977 GACGAGATTGAGATTGAGATTGAGATTGAGATTGAGATTGAGATTGAGATTGAG 2035
Qy 670 LeuCysHisLeuLeuPheAlaGluGluGluGluGluGluGluGluGluGluGlu 680
Db 2036 ATATGTCAAGCTTGTTTACAAAAAAGCAGAG 2068

RESULT 9
HUMORFX LOCUS HUMORFX 3028 bp mRNA linear PRI 06-OCT-2001
DEFINITION Human mRNA for KIAA0043 gene, complete cds.
ACCESSION D26362
VERSION D26362.1 GI:452518
KEYWORDS KIAA0043.
SOURCE Homo sapiens male myeloblast cell_line KG-1 cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nomura, N., Nagase, T., Miyajima, N., Saruka, T., Tanaka, A., Sato, S.,
Seki, N., Kawarabayashi, Y., Ishikawa, K. and Tabe, S.
TITLE Prediction of the coding sequences of unidentified human genes. II.
The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
analysis of cDNA clones from human cell line KG-1
JOURNAL DNA Res. 1 (5), 223-229 (1994)
MEDLINE 96051398
REFERENCE 2 (bases 1 to 3028)
AUTHORS Ohara, O., Nagase, T., Kikuno, R. and Nomura, N.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1993) Osamu Ohara, Kazusa DNA Research Institute;

```

```

FEATURES
source
    1532-3, Yana, Kiearazu, Chiba 292-0812, Japan
    (E-mail:cdna1nfo@kazusa.or.jp, Tel:+81-438-52-3913)
    Location/Qualifiers
        1..3028
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /sex="male"
        /cell_line="KG-1"
        /cell_type="myeloblast"
        1..3028
        /gene="KIAA0043"
        1..139
        /gene="KIAA0043"
        5' UTR
        140..2320
        CDS
        /gene="KIAA0043"
        /note="similar to Human homolog of Drosophila female
        sterile homeotic mRNA ( HUMSHG )"
        /codon_start=1
        /protein_id="BAA05393.1"
        /db_xref="GI:452519"
        /translation="MSTATTVAPAGIPATPGVPNPPEVSNPKRKRKTNOLOYMON
        VVKTLMGHOFAMPYOPVDALIKLIPYHKIKRPMMGTKRLENNYVWASSEM
        ODFTMTFNHCYIYNKPTDDIYLMQALEKIKQVAKQPOBEVELLPAPKQKRPKPA
        AGAQSAGTQVVAANSVSPATPFQSVPTVSOTPIATPVPTTANVTSVPVPAVA
        PPPATPIVPVPPPTPVKKGVKRAKADTTPTTSATITASSESPPLSDPKQAKV
        ARRESGGRPIKPKKDLDEGEVPOHAGKKLSBELRYCDILREMLSKHAAVAMP
        KRPVDALELDHYDILIKHPMDLSITVKRMDGREYPPAOGFPAADVRLMFSCYKYP
        PDHEVVAARLQDVFEMRFAPKMPDEYVPAALPAPAPMSVKAASSRSSSESSDS
        GSSDSEERATRLALQEOQLAVHEQQLAALSOAPVNRKKKKKKKKKKKKKKKKKK
        EKHVKAEERKAKVAPPAKOQKQKAPAKKANSTTTGROLRKQKQKASASVDSSE
        BEGLPMSVDERKQLSLDINRLPGLKGLGVNHIIGREPSLRNSNDELEIPEITLKP
        TLRELERYVKSCLOQKORRPFASGSKQKASKEBLAEKKELEKRLQDVSGQLSS
        KKPARKRGSPSGGPERLSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
        2321..3028
        3' UTR
        /gene="KIAA0043"
        BASE COUNT 758 a 864 c 877 g 529 t
        ORIGIN

Alignment Scores:
Pred. No.: 1,39e-70 Length: 3028
Score: 1922.00 Matches: 403
Percent Similarity: 69.51% Conservative: 69
Best Local Similarity: 59.35% Mismatches: 151
Query Match: 52.82% Indels: 57
DB: Gaps: 9 Gaps: 16

US-09-700-590a-22 (1-688) x HUMORFX (1-3028)
Qy 25 GlnMetSerThryGlnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThr 43
Db 137 AGGATGTCACCGCCGACGACGACGCGCCCGGGGAGATCCCGGACCCCGGCCCTGTG 196
Qy 44 AsnProProProProGluThrySerAsnProAlaLysProLysArgGlnThryAsnGlnLeu 63
Db 197 AACCCACCCCCCGGAGAGGTCTCCAAACCCACAGAGCCCGGCGGCAAGCAACAGCTG 256
Qy 64 GlnThryLeuLeuArgValValLeuLysThryLysThryHisGlnPheAlaTrpProPhe 83
Db 257 CAGTACATCAGAAATGTGTGTGTGAAGACGCTGTGAAACACACAGTTCCCTGGCCCTTC 316
Qy 84 GlnGlnProValAlaPheAlaValLysLeuAsnLeuProAspThryThryLysIleLeuLysThr 103
Db 317 TACCAAGCCCGTGAAGCCATCAAAATTTGAACCTCGCGGATTATCATTAATTTAAAAAC 376
Qy 104 PrometLysPheGlyThryLysLysLysArgLeuGluAsnAsnThryThryTrpAsnAlaGln 123
Db 377 CCAATGATATGGGGGACTATTAAAGAAAGAACTAAGAAATTAATTAATTAATTAATTA 436
Qy 124 GluCysIleGluAspPheAsnThryMetPheThraAsnCysThryLysThryLysProGly 143
Db 437 GAATGTATCAGAGACTTCAACCACTGTTTACAAATTTTATTAACAAGCCACACA 496
Qy 144 AspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLysIleAsnGln 163

```



```

Db 497 GATGACATAGTGTATGCGCCAGCTTTAGAGAAATTTTCTACAAAAGTGGCCAG 556
Qy 164 LeuProthrgluGluThrgluIleMetIleValGlnAlaIleSgIlyArgIlyArg----- 181
Db 557 ATGCCCCAAGAGAGAGTGAATTATATACCCCTGCTCCAAAAGGGCGAAGCCG 616
Qy 182 -----GlyArgIlySgIlyThrgluThrAlaIlyProGlyValSerThrValProAsnThr 199
Db 617 GTCGCGGAGCCCGACGAGCGGATCACAGCA-----GTGGGGCGCGTGTCTCTGTC 670
Qy 200 ThrGlnAlaSerThrProProGlnThrglnThrPro---GlnProAsnProProProVal 218
Db 671 TCCCCACGCG-----ACCCCTTTCCAGACGTCGCCCCCAACCGCTC 709
Qy 219 GlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProVal 238
Db 710 TCCCAAGCGCC-----GTCAATCGCTGCCACCCCTGTATA 742
Qy 239 MetThrValValProProGlnProLeuGlnThrProProProValProProGlnProGln 258
Db 743 CCAACCACTACGTCAAAC-----GTCACTGCTGCTCCACATCCCCCAAGTCCGCC 793
Qy 259 ProProProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaIleThr 278
Db 794 CCACTCTCT-----CTGCGCACCCCATCGTCCCGTGTCTCTCTACAG 838
Qy 279 ProGlnProValIlyThrgluIlyGlyValIlySargIlySalaAspThrThrThrProThr 298
Db 839 CCG-----CCTGTGCTCAAGAAAAGGCGTGAAGCGGAAGACACAAACCATCTCCACG 895
Qy 299 -----ThrIleAspProIleHisGluProProSerLeuProProGluProIly 314
Db 896 ACGTCGCGCATCACTGCCAGCGCGAGTAGTCGCCCGCCCGCTG---TCAGACCCCAAG 952
Qy 315 ThrThrIlySleuGlyGlnArgArgIlySer---SerArgProValIlyProProIlySls 333
Db 953 CAGCGCAAGTGTGTGCGCGCGGAGAGTGTGTGCGCGCCCATCAAGCTCCCAAGAG 1012
Qy 334 AspValProAspSerGlnGlnHisProAlaProGluIlySerSerIlyValSerGluGln 353
Db 1013 GACCTGAGAGCGCGAGGTGCCCGACAGCCAGCAAGCAAGGCGCAACTCTCGAGAC 1072
Qy 354 LeuIlyCysCysSerGlyIleLeuIlySgIlyMetPheAlaIlySlyHisAlaIleAla 373
Db 1073 CTGGCGTAATCGACAGATCTCTAGGAGATGCTATCAAGAAACAGCGCGCTACGCC 1132
Qy 374 TrpProPheIlyIlyProValAspValGlnAlaIleuGlyLeuHisAspTyrCysAspIle 393
Db 1133 TGCCCTTCTCAAGCCAGTGTGATGCCAGGCCCTGAGCTGCAAGTACCAAGACATC 1192
Qy 394 IleIlyHisProMetAspMetSerThrIleIlySerIlySleuGlnAlaArgGluTyrArg 413
Db 1193 ATCAAGACACCGATGAGCTCGACGACCTGTAAAGAAAGATGAGCGCGAGATACCCA 1252
Qy 414 AspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnCysTyrIlySerAsn 433
Db 1253 GACGACAGGGCTTGTCTGATGATCCGGCTGATGTTCTGAATTGCTTACAAATACAT 1312
Qy 434 ProProAspHisGluValAlaValAlaMetAlaArgIlySleuGlnAspValPheGluMetArg 453
Db 1313 CCCCCAGACCAAGAGTGTGGCCATGGCCCGGAAGCTCCAGAGACTGTGTTGAGATGAGG 1372
Qy 454 PheAlaIlyMetProAspGluProGluGluProValAlaValAlaSerSerProAlaVal 473
Db 1373 TTTGCCAAGATGCCAGATGAGCC-----GTGAGAGCACCGCGCGCTG 1414
Qy 474 ProProProThrllyValAlaAlaProProSerSerSerAspSerSerSerAspSer 493
Db 1415 CTTCCCGCGCGCGCCCATGTGTGACAAAGCGCTGAGAGCAGCGCTGACGTAGAGAG 1474
Qy 494 SerAspSerAspSerSerThrAspAspSerGluGluArgAlaGlnArgLeuAlaGlu 513

```

```

Db 1475 AGCTCTTCGACTCAGGAGCTCTCGAAGCTCGAGAGAGCGGCGCACAGCTGGCGAG 1534
Qy 514 LeuGlnGluIleuIlySalaValHisGluGlnLeuAlaIleuSerGlnProGlnGln 533
Db 1535 CTGCAGAGAGAGCTGAAGGCGGTGACAGAGACAGCTGCGCGCTGTCTAGGCCCCAGTA 1594
Qy 534 AsnIlyProIlySls-----LysGluIlyAspIlySlySgIlyIlySls 547
Db 1595 AACAAACCAAGAAAGAAAGAGAAAGAGAGAGAGAGAGAAAGAAAGAGACAAAGAG 1654
Qy 548 LysGluIlyHisIlySargIlySgIlyGluValGluGlnAlaAsnIlySlySerIlySls 564
Db 1655 AAGAGAGAGAGAGAGCAAGTGAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1714
Qy 565 ---AlaIlySgIlyProProIlySlySlyThrllySlyAsnAsnSerSerAsnVal 583
Db 1715 CTGCGCAAGAGAGCTCAGCAAGAAAGAGCTCTGCCAAGAGGCCAAACAGACAGCAG 1774
Qy 584 SerIlySgIlyProAlaProMetIlySerIlySerIlyProProThrllyGluSgIlyGlu 603
Db 1775 GCGGAGAGAGAGCTGAAGAAAGCGCGCAAGCAGCATCTGCTCTACGACTCAGAGGAA 1834
Qy 604 GluAspIlyCysIlyProMetSerIlyGluGluIlyArgGlnIleuSerLeuAspIleAsn 623
Db 1835 GAGAGAGAGAGAGCTGCGCATGAGCTAGTGAAGAGCGCGAGCTTGAAGCTGACATCAAC 1894
Qy 624 LysIleuProGlyGluIlySleuGlyArgValAlaHisIleIleGlnSerArgIlyProSer 643
Db 1895 CGGCTGCCCGGAGAGAGTGTGGCGGGTGTGTCACATCTCCATCTCGGAGCGCTCG 1954
Qy 644 LeuIlyAsnSerAsnProAspGluIleGluIleAspPheGluThrllySlyProSerThr 663
Db 1955 CTCAGGAGACTCAACCCCGCAGAGTGAATGACTTGAAGCTGTGAAGAACCCACACT 2014
Qy 664 LeuArgIlyLeuGlyAlaLeuCysHisLeuLeuPheAlaGluIly---LysGluThr 681
Db 2015 TTGGGGGAAGCT-GGAGAGATATGCAAGTCTTGTACAGAAAAGCAAGAGAAACC 2070

RESULT 10
AF045462 3649 bp mRNA linear ROD 01-JUL-1998
LOCUS AF045462
DEFINITION Mus musculus female sterile homeotic-related protein Frig-1 mRNA,
complete cds.
ACCESSION AF045462
VERSION AF045462.1 GI:3273700
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3649)
AUTHORS Rhee, K., Brunori, M., Besset, V. and Wolgemuth, D.J.
TITLE Expression and potential role of Frig-1, a putative murine
bromodomain-containing homologue of the Drosophila gene female
sterile homeotic
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3649)
AUTHORS Rhee, K., Brunori, M., Besset, V. and Wolgemuth, D.J.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-1998) Department of Genetics and Development,
Columbia University College of Physicians and Surgeons, 630 W 168th
St., Black Bldg 163, New York, NY 10032, USA
location/Qualifiers
FEATURES
1..3649
/organism="Mus musculus"
/db_xref="taxon:10090"
681..3077
/note="bromodomain-containing protein"
/codon_start=1
/product="female sterile homeotic-related protein Frig-1"
/protein_id="AAC24810.1"
/db_xref="GI:3273701"
/translation="MLQNTVPHKLPEGNAGLGLGPAAAPKRIKPSLLYGFES

```



```

PTMASVPLQLAPNPPEVSNPKKPGVTVNQLQYHKVNMALMKEHPAMPFPROV
DAVKGLPDYHKI IKOPMDGTI KRLENNYMAWSECMQDPTMTNICYINPKPD
IYMAQTELEI FLQKVASMOEOLAVTI PKNSHKSGKLAQSLTSAHPAVS
SVSHATLYPPEIPTTVNIIPSVISLPLKLSHAPPLVSAAPADPLAKK
GVKRAADTPPTPTIALARGSPASPGSLPKRAALPMRRESGPRKPRDLPDPO
QOHOSKKKLSBOLKNCNGILKELSKHAAVMPFYRPVDSALGLHDYDIIKH
MLSTVKKRMENRVDQEPADVRLMNCYNP PHDVAMARKLDVFEERYA
KMPDELEGPPLVSTALPPGLTKSSSESSSESSSESSSESSSESSSESS
DSEERAHRLAQEQLRAVHEQLALSGPISKPRKREKKEKKKRAKHGRI
IDEDGPRAPRPPPOPKSKKAGGGGNNTTLSHPFGSGGSKL PKSGDAPV
LPGYDSESESRSPMSYDEKRLSDIKLGEKGRVYHIIQAREPSRPSBEI
EIDFTLKPSITREIRYVLSCKRKPRPYITRKVGKTKBELAEKRELEKLD
VSGQNSTKPKPKASEKTESSAQVLVPSRLAASSSSSDSSSSSSSDTSDBSG
"
BASE COUNT      920 a      986 c      976 g      767 t
ORIGIN
Alignment Scores:
Score:           1,2e-69      Length:      3649
Percent Similarity: 66.94%      Matches:      403
Best Local Similarity: 55.51%      Conservative: 83
Query Match:      52.24%      Mismatches: 145
DB:               10          Indels:      96
                       Gaps:         18
US-09-700-590A-22 (1-688) x AF045462 (1-3649)
QY      5 SerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGluThrSer 24
DB      756 GCACGACCCGAGAAAGGATTCGAAAGCCTTCTGCTGTATGAGGATTTGAGAGCCCC 815
QY      25 GlnMetSerThrThrglnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThrAsn 44
DB      816 ACAATGGCTTCTGTACAGCT-----TTACAACGCGCCCTCGCCCAAT 857
QY      45 ProProProProGluThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGln 64
DB      858 CCAACCAACCCCTGAGGTGTCACCAAAAGCCAGGAGCGGGTACCAAAACCAACGTCAG 917
QY      65 TyrLeuLeuArgValValLeuLysThrLeuTyrLysHisGlnPheAlaTrrProPheGln 84
DB      918 TACCTGACAGGATGATGATGAGGCTGTGTGAGGATGATGATGATGATGATGATGATGATG 977
QY      85 GlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleIleLysThrPro 104
DB      978 CAGCCTGTGACCGCTGTGAAGCTGTGCGGATTTATACAAATTAATAAACAAGCCT 1037
QY      105 MetAspMetGlyThrIleLysLysArgLeuGluAsnAsnTyrTyrTrrPasnAlaGlnGlu 124
DB      1038 ATGACATGAGGATGATGATGAGGAGCTGTGAAACCAATTACTACTAGGCTGCTCGAGAA 1097
QY      125 CysIleGlnAspPheAsnThrMetPheThrAsnGlyTyrIleTyrAsnLysProGlyAsp 144
DB      1098 TGTATTCAGAGATTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1157
QY      145 AspIleValLeuMetAlaGlnAlaLeuGluLysLeuPheLeuGlnLysIleAsnGluLeu 164
DB      1158 GATATTGTCTTATGACACAGACATCGAATAATCTTCTTACAGAAAGTGCATCATG 1217
QY      165 ProThrGluGluThrGluIleMetIleValGlnAlaLysGly-----ArgGly--- 180
DB      1218 CCAAGAGAGGAGCAAGAGCTGTGTGATCCATCCCTAAAAACAGCCATAAGAGGGGGG 1277
QY      181 -----ArgGlyArgGlyGluThrGlyThrAlaLysProGlyValSerThr 195
DB      1278 AAGTAGCAGACCTCGAGGAGCATATTACAGGCCATCAGGTGCTGTCTTCTTCT 1337
QY      196 ValProAsnThrThrglnAlaSerThrProGlnThrGlnThrProGlnProAsnPro 215
DB      1338 GTGTCCGATACAGCCCTGTATACACCACTCGAATA-----CCT 1379
QY      216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235

```


D	1778	TCGCACTTCGAAGCTCTTGAC	-----TCGCTGGACCCCGGCTCTT	1825
Q	256	GlnProGlnProProAlaProAlaProGlnProValGlnSerHisProProlIle	275	
D	1826	GTTACTGACAGCTCTCCAGCC	-----1846	
Q	276	AlaAlaThrProGlnProValLysThrLysGluValLysArgLysAlaSerThr	295	
D	1847	-----CAGCCCTTGGCCAAAGAAAAGGGGTAAAGCGGAAGCAATCTAC	1894	
Q	296	ThrProThrThrIle-----AspProIleHisGlnProProlSerLeu	309	
D	1895	ACCCCTACACTACAGCCATCTGGCTCGCTGTTCCAGCTACAGCCCTCTGGAGCTT	1954	
Q	310	ProProGlnProLysThrLysLeu--GlyGlnArgGluSerSerArgProVal	328	
D	1955	-----GAGCCTAAGCGACAGACGGCTTCCCGCTATCGTAGAGAGATGTGGCCCATC	2008	
Q	329	LysProProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSer	348	
D	2009	AAGCCCAAGCAAGCAAGCTTGGCTGTGCTGACCAACACACAGAGCTTAAGAAAGA	2068	
Q	349	LysValSerGlnGlnLeuLysCysCysSerGlyIleLeuLysGluMetPheAlaLysLys	368	
D	2069	AAGCTTTCAGAACAGTTAAACATTCATTCGATTTGAAGAGTACTCTTACAG	2128	
Q	369	HisAlaAlaThrAlaTrpProPheTyrLysProValAspValGluAlaLeuGlyLeuHis	388	
D	2129	CATGCTCGTATGCTGGCTCTTCTATTAACAGTGATGCTTTCGACCTGGCTGCAT	2188	
Q	389	AspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGlu	408	
D	2189	GACTACCATGACATCTTAAAGCAACCCATGAGACTCAGACATGTAACGGAAAGTATG	2248	
Q	409	AlaArgGlnTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsn	428	
D	2249	AACCGTATTAACCGGATGACACAGAGTTTGCTGCTGATGTACGGCTTATGTTCTCCAC	2308	
Q	429	CysTyrLysTyrAsnProProAspHisGluValAlaMetAlaArgLysLeuGlnAsp	448	
D	2309	TGCTATTAAGTACATCCCCCAGATCAAGATGTTGTGGCAATGGACGAAGCTACAGAT	2368	
Q	449	ValPheGluMetArgPheAlaLysMetProAspGluProGluGluProValAlaVal	468	
D	2369	GTAATTGACTCGTATGATGCAAGATGCAAGTGAACCACTGAACCAAGGCGCTTACCA	2428	
Q	469	SerSerProAlaValProProPro--ThrLysValAlaLProProlSerSerSerAsp	487	
D	2429	GTCCTTACTCCCAATGCCCCCTGGCTGGCCAAATGCTCTTCAGAGCTCTCCAGTAGGAA	2488	
Q	488	SerSerSerAspSerSerSer-----Asp	495	
D	2489	AGTAGCATGATGACCTCTTGAGAGAGAGAGAGAAAGATGAGAGAGACAGAGAGAA	2548	
Q	496	SerAspSerSerThrAspAspSerGlnGluArgAlaGlnArgLeuAlaGluLeuGln	515	
D	2549	GAAAGAGTGAAGCTCAGACTCAGAGAGAAAGGGCTCATCGCTTAGCAGAACTACAG	2608	
Q	516	GluGlnLeuLysAlaValHisGluGlnLeuAlaAlaLeuSerGlnProGlnIleAsnLys	535	
D	2609	GAAACAGCTTGGGACATGACATGAACAATCGCGCTCTGTCCCAAGGTCATATACAAAG	2668	
Q	536	ProLysLysLysGluLysAspLysGluLysLysLys-----GluLysHis	551	
D	2669	CCCAAGAGGAAA--AGAGGAAAAAGAAAAAGAAAGAAACGGAAGGCAGAGAGAT	2725	
Q	552	LysArgLysGluGluValGluGluLeuLysSer--LysAlaLysGluProPro--	569	
D	2726	CGAAGCGAGCTGGGGCCGACGAAGATGACAAAGGGCGCTAAGGACACCCGCCCACTCA	2785	
Q	570	ProLysLysThrLysLysAsnAsnSerSerAsnValSerLysLysGluProAla	589	
D	2786	CTTAAGAAGTTCAGAAAGCAAGATGGCAGTGGGGGTGACAGTGTCTTAGAGCCCTTCT	2845	

QY	590	-----PromElyseirys 594
Db	2846	GGCTTTGACCTTCTGGAGAGAGTGGACCAAGCTCCCAAAAAAGGCCAACAAAGACACC 2905
QY	595	ProPro-----Prothr---TyrGluserIguIguIuAplysCysIysProMetSer 611
Db	2906	CCACCTGGCCCTGCTACAGGTTATGATTGATCAGAGGAGGAGGAAGACGACCCATGAGT 2965
QY	612	TyrIguIguIysAArgIInleuSerIleuAspIleuInlyIeuProGlyGluIlyIeuGly 631
Db	2966	TACATGATGAAGCGGACGAGCTGAGCTGACATCAACAATTACCTGGGGAAGAGCTGGGC 3025
QY	632	ArgValValHisIleIleGlnSerArgIuIuProSerIleuIysAsnSerAspProAspIu 651
Db	3026	CGAGTTGGCATATATATCCAGCCAGGAGGCCCTCTTACGTGATTCAACCCAGAACAG 3085
QY	652	IlegIuIeAspPheGluThrIleuIysProSerThrIleuAArgIuIeGluIyAlaIeuCys 671
Db	3086	ATTGAGATTGATTTTGAACAAGCTCAAGCATTCACTTAGAGAGCTTGA-GCGCTATGT 3144
QY	672	HisIeuIeu 674
Db	3145	CCTTTCCTGG 3153
RESULT 12		
LOCUS	HUMFSHG	4053 bp mRNA linear PRI 22-JUN-1995
DEFINITION	HUMFSHG	Human homolog of Drosophila female sterile homeotic mRNA, complete cds.
ACCESSION	M80613	
VERSION	M80613.1	GI:182768
KEYWORDS	MHC class II antigen; female sterile homeotic gene.	
SOURCE	Homo sapiens cDNA to mRNA.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 4053)	
TITLE	Beck, S., Hanson, I.M., Kelly, A., Pappin, D.J.C. and Trowsdale, J.	
JOURNAL	A homologue of the Drosophila female sterile homeotic (FSH) gene in the class II region of the human MHC	
REFERENCE	DNA Seq. (1991) In press	
AUTHORS	2 (sites)	
	Haynes, S.R., Dollard, C., Winston, F., Beck, S., Trowsdale, J. and David, I. B.	
TITLE	The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins	
JOURNAL	Nucleic Acids Res. 20 (10), 2603 (1992)	
MEDLINE	92285152	
FEATURES	1350857	
source	Location/Qualifiers	
	1..4053	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/cell_type="T CELL"	
	1178...3442	
	/note="putative"	
	/codon_start=1	
	/protein_id="AAA68890.1"	
	/db_xref="GI:182769"	
	/translation="MASVPALQLTPANPPPEVSNPKKGRVTNQLYLHKVYWKALM; KHPFAMPFQPDVAVKLGLPDYHKIIKQPMDSKTKRLNNYMAESGODPNTMFGTNCYINVKRPDDIVLMAQTELEKIFLOVAMPQEOELAVTI PINSKKGAKLAALQGLDHYDIIKHPMDSLTVKKRMENRDYRQAEPAADVRLFNSCYKNPDDHVVANA RKIQDVEEFYAKMPDEPLPGPLPVSTAMPQGLAKSSSSSSSSSSSESEEEEEEE DEDEDEESSSESSSEERHARHLAEIOELAVHEOLASOGPI SKPKRREKKEKKKKR KKKAKHGRGACADDDGPRAPRPPQPKSKKASGGGSAIAGSGPGSGSGGT KLPRKATKTPRALPFGCYDSEEESESPSYDKRQSLDINTLKPGKLGARVYIIQA RESSLNPPSEELIEDPTELTPESTLELELYVLSCKLKKRRKPTTKPVGKTKBELA LEKKRELEKRLQVDSQGLNSTKPKPKANKETSSSKQYAVARLSKSSSDSSSSSS	
CDS		

SSSSSDPDSG
polya_signal
4015.4020
/note="putative"
BASE COUNT 997 a 1132 c 1067 g 857 t
ORIGIN

Alignment Scores:
Pred. No.: 4.48e-69 Length: 4053
Score: 1888.00 Matches: 404
Percent Similarity: 67.08% Conservatave: 81
Best Local Similarity: 55.88% Mismatches: 142
Query Match: 51.88% Indels: 97
Gaps: 20

US-09-700-590a-22 (1-688) x HUMFSHG (1-4053)

QY 5 SerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGlnThrSer 24
Db 1115 GCACACACGAGGAAAGATTGCAAAACCTCTCTGTATGAGGCTTTGAGAGCC 1174
QY 25 GlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThrAsn 44
Db 1175 ACAATGGCTTGCGCTGCTCTTGCATCTTACCCCTGCC-----AAC 1216
QY 45 ProProProProGlnThrSerAsnProAsnProLysProLysArgGlnThrAsnGlnLeuGln 64
Db 1217 CCACACACCCCGGAGGCTGCTCAATCCCAAAAGCCAGGACGAGTACCAACAGCTGCA 1276
QY 65 TyrLeuLeuArgValValLeuLysThrLeuTyrPlySHISGlnPheAlaTrpProPheGln 84
Db 1277 TACCTACCAAGTAGATGATGAGGCTCTGTGAAACATCATGTCGATGGCATTTCCGG 1336
QY 85 GlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleIleLysThrPro 104
Db 1337 CAGCGTGTGATGCTGTCAAACTGGGCTTACCGGATTTACAAATTTATTAACAGCT 1396
QY 105 MetAspMetGlyThrIleLysLysArgLeuGlnAsnAsnTyrTyrTrpAsnAlaGlnL 124
Db 1397 ATGGACATGGGACTATTAAAGAGACTTGAAACAAATTATTATGGGCTGCTCAGAG 1456
QY 125 CysIleGlnAspPheAsnThrMetPheThrAsnGlyTyrIleTyrAsnLysProGlyLysP 144
Db 1457 TGTATGCAAGATTTTAAATACATGTTCAACAATCTTACATTTAACAAAGCCCAT 1516
QY 145 AspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGlnLeu 164
Db 1517 GATATTGCTTAATGACCAACGCTGGAAAGATATTCCTACAGAGGTTGCATCAATG 1576
QY 165 ProThrGlnGlnThrGlnLysIleMetIleValGlnAlaLysGly-----ArgGly--- 180
Db 1577 CCACAGAGAAACAAAGAGCTGTGTAGTACCATCCTTAAGAACAGCCACAAAGAGGGCC 1636
QY 181 -----ArgGlyArgLysGlnThrGlyThrAlaLysProGlyValSerThr 195
Db 1637 AAGTTGACGAGGCTCCAGGCGCAGTGTACAGTGCACAGTCCCTCCCTCTCTCT 1696
QY 196 ValProAsnThrThrGlnAlaSerThrProProGlnThrGlnThrProGlnProAsnPro 215
Db 1697 GTGTACACACAGCCCTGTATATCTCTCCACTGAGAT-----CCT 1738
QY 216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235
Db 1739 ACCACTGCTTCAACATCCCAACCA-----TCAGTCAATTCC 1777
QY 236 ThrProValMetThrValValProProGlnProLeuGlnThrProProProValProPro 255
Db 1778 TCTCCACTTCAAGTCTCTTGAC-----TCGTGACAGCCCGCTCTCTGCT 1825
QY 256 GlnProGlnProProProProAlaProAlaProGlnProValGlnSerHisProProIleIle 275
Db 1826 GTTACTGACGCTCTCCAGCC----- 1846
QY 276 AlaAlaThrProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThr 295

Db 1847 -----CAGCCCTTCCCAAGAAAAAGGCTTAAGCGGAAAGAGATCTAAC 1894
QY 296 ThrProThrThrIle-----AspProIleHisGlnProProSerLeu 309
Db 1895 ACCCTTACACTACAGCATCTTGCTCTGTTCTCCAGCTAGCCCTCTGGAGTCTT 1954
QY 310 ProProGlnProLysThrThrLysLeu--GlyGlnArgArgLysSerArgProVal 328
Db 1955 -----GAGCTTAAGGACAGCGCTTCCCTTATGCGTAGAGAGATGTGCGCCCATC 2008
QY 329 LysProProLysLysAspValProAspSerGlnGlnHisProAlaProGlnLysSerSer 348
Db 2009 AAGCCCCGCAAGACCTGCTGACTCTCAGCAACACACAGAGCTCTTAACAAAGGA 2068
QY 349 LysValSerGlnGlnLeuLysCysCysSerGlyIleLeuLysGlnMetPheAlaLysLys 368
Db 2069 AAGCTTCAAGACGTTAAACATTCGAATGCAATTTGAAAGAGTTACTCTTAAGAG 2128
QY 369 HisAlaAlaTyrAlaTrpProPheTyrLysProValAspValGlnAlaLeuGlyLeuHis 388
Db 2129 CATGCTGCTATGCTTGCGCTTCTTATTAACAGTGTATGCTTGCACATTGGCTGCAT 2188
QY 389 AspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGln 408
Db 2189 GACTACCATGACATCATTAAGCACCCCATGAGCCTCAGCACTGTCAGCGCAAGATGAG 2248
QY 409 AlaArgGlnTyrArgAspAlaGlnGlnPheGlyValAspValArgLysMetPheSerAsn 428
Db 2249 AACCCTGATTCGCGGATGCAACAGAGTTTGCTCTGATGTACGGCTTATGTTCTCCAC 2308
QY 429 CysTyrLysTyrAsnProProAspHisGlnValAlaAlaMetAlaArgLysLeuGlnAsp 448
Db 2309 TGCTATTAAGTACATCCCGGATCAGATGTCGATGTGTGCAATGGCAGCAAGACTACAGAT 2368
QY 449 ValPheGlnMetArgPheAlaLysMetProAspGlnProGlnGlnProValAlaVal 468
Db 2369 GTATTGAGTTCCTTATGCAAGATGCCAGATGACACTGAACCAAGGGCTTTTACA 2428
QY 469 SerSerProAlaValProProPro--ThrLysValAlaAlaProProSerSerSerAsp 487
Db 2429 GTCTTACTGCTGATGCCCTCGCTGCTGGCCAAATCGTCTTCAAGTCTCTCAGTAGAGAA 2488
QY 488 SerSerSerAspSerSerSer-----Asp 495
Db 2489 AGTAGAGTAGAGACTCTCTGAGCAAGAGAGAGAGATGAGAGAGAGAGAGAGAA 2548
QY 496 SerAspSerSerThrAspAspSerGlnGlnLysArgAlaGlnArgLysGlnLeuGln 515
Db 2549 GAAGAGAGTGAACCTCAGACTCAGAGAGAGAGAGAGGCTCATCGCTTAAGCAACTACAG 2608
QY 516 GlnGlnLeuLysAlaValHisGlnGlnLeuAlaLeuSerGlnProGlnGlnAsnLys 535
Db 2609 GAACAGCTTCGGGAGTAGACATGAACAATGCTGCTCTGTGCCAGGGTCCAAATTCACAG 2668
QY 536 ProLysLysLysGlnLysAspLysLysGlnLysLysLys-----GlnLysHis 551
Db 2669 CCACAGAGAGAA--AGAGAGAAAAAGAGAAAAAGAGAAACGAGAGGACGAGAGACAT 2725
QY 552 LysArgLysGlnGlnLysAlaGlnGlnAsnLysSer--LysAlaLysGlnProPro-- 569
Db 2726 CGAGGCGCAGCTGGGGCGGAGAGAGATGACAAAGGGCTTAGGGGACCCCGCCACCTCAA 2785
QY 570 ProLysLysThrLysLysAsnAsnSerSerAsnSerAsnValSerLysLysGlnProAla 589
Db 2786 CTTAAGAGATCAAGAAAGCAAGTGGCGTGGGGCTGCACTGCTCTTAGGCTCTTCT 2845
QY 590 -----PrometLysSerLys 594
Db 2846 GGTCTTGACCTTCTGAGAGAGTGGACCAAGCTCCCAAAAAGGCGCAAAAGACAGGCC 2905
QY 595 ProPro-----ProThr--TyrGlnSerGlnGlnLysCysLysPheMetSer 611


```

Db 3034 GTATTGATGTCCTTATGCGAAGATGCGACATGAAACCACTAGAACAGGCGCTTTACCA 3093
Qy 469 SerSerProAlaValProProPro---ThrIysValAlaProSerSerSerAsp 487
Db 3094 GTCCTACTGCGAAGCCCCCTGGCTGGCCAAATCGCTTCAAGACTCTCCAGTGAGGAA 3153
Qy 488 SerSerSerAspSerSerSer-----Asp 495
Db 3154 AGTAGCAGTAGAGACTCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 3213
Qy 496 SerAspSerSerThrAspAspSerGluGluGluAlaGlnAlaGluLeuGln 515
Db 3214 GAAGAGAGTAGAAGCTCAAGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 3273
Qy 516 GluGluLeuValAlaValIleGluGluLeuAlaAlaLeuSerGlnProGluGlnAlaAsnLys 535
Db 3274 GAACAGCTTCGGGAGTAGACATGAAACACTGGCTGCTCTGTCTCCAGGCTCAATATCCAA 3333
Qy 536 ProLysLysLysGluLysAspLysLysGluLysLysLys-----GluLysHis 551
Db 3334 CCAAGAGAGAGAA---AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCAT 3390
Qy 552 LysAlrLysGluGluValGluGluLysLysSer---LysAlaLysGluProPro--- 569
Db 3391 CGAGGCGGAGCTGGGGCCGATGAAGATGACAAAGGGGCGCTTACGAGGCCCGCCCACTCAA 3450
Qy 570 ProLysLysThrLysLysAsnAsnSerSerAsnSerValSerLysLysGluProAla 589
Db 3451 CCTAAGAGAGTCACAGAAAGCAAGTGCGAGTGGGGGTGCGACAGTGTCTTTAGGCCCTTCT 3510
Qy 590 -----PrometLysSerLys 594
Db 3511 GGCCTTGGACCTTCTGAGAGAGAGTGACCAAGCTCCCAAAAAGGCCAACAAAGACAGCC 3570
Qy 595 ProPro-----ProThr---TyrGluSerGluGluLysLysCysLysProMetSer 611
Db 3571 CCACCTGCGCTGCTCAAGCTTATGATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3630
Qy 612 TyrGluGluLysArgLysLysLeuSerLysLeuAsnLysLeuProGluGluLysLeuGly 631
Db 3631 TACAGATGAGAGAGCGGACCTGAGCTGAGCATCAACAAATTAACCTGGAGAGAGAGAGAGAG 3690
Qy 632 ArgValAlaHisIleIleGlnSerArgLysProSerLysAsnSerAspProAspGlu 651
Db 3691 CGAGTGGCATATATATCAAGCCAGGAGGCCCTTTACGATGATCAACCCAGAGAGAG 3750
Qy 652 IleGluIleAspPheGlnThrLeuLysProSerThrLeuAlaGluLeuGlyAlaLeuCys 671
Db 3751 ATTGAGATTGATTGTAACACTCAAGCCATCCACATTAGAGAGCTTGA--GCGCTATGT 3809
Qy 672 HisLeuLeu 674
Db 3810 CCTTCTCTG 3818

RESULT 14
HUMKIA9001 4664 bp mRNA linear PRI 09-FEB-1999
LOCUS HUMKIA9001 4664 bp mRNA linear PRI 09-FEB-1999
DEFINITION Human mRNA for KIAA9001 gene, complete cds.
ACCESSION D42040
VERSION D42040.1 GI:577292
KEYWORDS KIAA9001.
SOURCE Homo sapiens male myeloblast cell_line KG-1 cDNA to mRNA.
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4664)
AUTHORS Nomura,N.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4664)
AUTHORS Nomura,N.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1994) Nobuo Nomura, Kazusa DNA Research
INSTITUTE, Gene Structure 1, 1532-3 Yana, Kisarazu, Chiba 292,

```

```

FEATURES
    source          Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp,
                    Tel:0438-52-3930, Fax:0438-52-3931)
                    Location/Qualifiers
                        1..4664
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /sex="male"
                        /cell_line="KG-1"
                        /cell_type="myeloblast"
                        1..4664
                        /gene="KIAA9001"
                        /gene="KIAA9001"
                        /gene="KIAA9001"
                        /note="female sterile homeotic (fsh) homolog RING3"
                        /codon_start=1
                        /protein_id="BA07641.1"
                        /db_xref="GI:577293"
                        /translation="MQONTVPHNKLDEGNAGILGIPAAAPGKRIKPSLLYEGFE
                        /SPMASVPAOLTPNPPEPEVSNPKPKPQVITNOQLHKKVVKALMKIOFAPWFPQ
                        VDAVKGLDPDYHKI IKQPMDMGTIKRLNNYVMASECMODENTFTNYCIYINRPTD
                        DIVAMQTLTEKIFLOKVASMPQEBELVVTIPKNSHKGAKLALGQVTSAAQVAV
                        SVSRNALVTPEPITPTVLANIPHSVLSPIKSHSGAPLLATVTAAPPOPLAKK
                        KGVKRAVDUTTPPTATILAGSPASPSGLSEPAALPPMKRESGRPIKPPRODLPS
                        QOQHOSKKKLESEQLKHONGILKELLSKHAAYVPFYKPVDAALGLDHYDITIKH
                        PMDLSTVKRMENRDYDQEFADYRLMFSNKYKPNPDHVDVAAARKLOVFEERY
                        AKMDPELEPGLPVSTAMPGLAKSSSESSSESSSESSSESSSESSSESSSESSSESS
                        SSDESERARHLAEIOELRAYEOLALISQGPISKPKRREKREKKRKAKHKG
                        AGADEDDKPRARPPQPKSKKASGSGSALGSGSGSGSGSTKL PKYATAP
                        PALPTGYDSEEBESRPMSTDEKQSLINKIPKRLGRVHIIOAREPSLDNPE
                        EIDIDETLKPSTLRELYVLSQLKKRKPPTTKRPVKTEBILALEKRELRKL
                        ODVSGQLNSTTKPKPKKANETKSSSQOVAVSRLSASSSSSSSSSSSSSSSDTS
                        DSG"
BASE COUNT      1084 a 1336 c 1293 g 951 t
ORIGIN
Alignment Scores:
Pred. No.:      5,11e-69      length: 4664
Score:          1888.00      Matches: 404
Percent Similarity: 67.08%      Conservative: 81
Best Local Similarity: 55.88%      Mismatches: 142
Query Match:    51.88%      Indels: 97
DB:              9      Gaps: 20

US-09-700-590A-22 (1-688) x HUMKIA9001 (1-4664)
Qy 5 SerGluProGluThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGluThrSer 24
Db 1780 GCAGACACGAGGAGAAAGAGATTGAAAACCCCTCTCTGTATGAGGAGGCTTTGAGAGCCC 1839
Qy 25 GlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThrAsn 44
Db 1840 ACAATGGCTTCGGTCCCTGCT---TTGCAACTTACCCCTGCG-----AAC 1881
Qy 45 ProProProProGluThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGln 64
Db 1882 CCACCACCCCGGAGAGGTCTCAAAATCCAAAAGCCAGAGAGAGATTACCAACAGCTGCA 1941
Qy 65 TyrLeuLeuArgValAlaLeuLysThrLeuTrpLysHisGlnPheAlaTrpProPheGln 84
Db 1942 TACCTACACAGTAGTAGATGATAAGGCTCTGTGAGAAACATCACTTCGATGGCCATTCGG 2001
Qy 85 GlnProValAspAlaValAlaLysLeuAsnLeuProAspTyrTyrLysIleIleLysThrPro 104
Db 2002 CAGCCTGTGATGCTGTCAAACTGGGTCTTACCGGATTATCACAAAATTTATTAACAGCCT 2061
Qy 105 MetAspMetGlyThrIleLysLysArgLeuGlnAsnLysTyrTyrTPAsnAlaGlnGlu 124
Db 2062 ATGACATGGGTACTATTTAAGAGAGAGACTTAAACATTAATTATTTGGGCTGCTTCAGAG 2121

```

QY 125 Cys11EglnAspPheAsnThrMetPheThrAsnCysTyr11eTyrAsnLysProGlyAsp 144
 ||:|||||
 Db 2122 TGTATGCAAGATTATTAATACCATGTTTACCAACAGTATTATTAACAAGCCACATGAT 2181
 QY 145 Asp11eValLeuMetAlaGluAlaLeuGluLysPheLeuGluLys11eAsnGluLeu 164
 ||:|||||
 Db 2182 GATATTGTCCTTAATGAGCAAAACCGTGAAGAAAGATATTCCTACAGAAAGTTGCATCAATG 2241
 QY 165 ProThrGluGluThrGluLeuMet11eValGlnAlaLysGly-----ArgGly--- 180
 ||:|||||
 Db 2242 CCACAGAAAGAAACAGAGCGTGTAGTACCATCCCTTAAGAAACGCCAACAGAAAGGGGGCC 2301
 QY 181 -----ArgGlyArGlyGluThrGlyThrAlaLysProGlyValSerThr 195
 ||:|||||
 Db 2302 AAGTTGGACAGCGCTCCAGGCGCATGTTACCATGCCATCAGGAGGCTGGCGTCTTCT 2361
 QY 196 ValProAsnThrThrGlnAlaSerThrProProGlnThrGlnThrProGlnProAsnPro 215
 ||:|||||
 Db 2362 GTGTCAACACAGCGCTGTATCTCTCCACCTGAGATTA-----CCT 2403
 QY 216 ProProValAlaAlaThrProHisProPheProAlaValThrProAspLeu11eValGln 235
 ||:|||||
 Db 2404 ACCACTGCTCTCAACATCTCCACCA-----TCAGTCATTCTCC 2442
 QY 236 ThrProValMetThrValValProProGlnProLeuGlnThrProProProValProPro 255
 ||:|||||
 Db 2443 TCTCACTCTCTCAAGCTTCGAC-----TCTGCTGAGACCCCGCTCTTGTCT 2490
 QY 256 GlnProGlnProProProAlaProAlaProGlnProValGlnSerHisProPro11e11e 275
 ||:|||||
 Db 2491 GTTACTGACAGCTCTCTCCAGC----- 2511
 QY 276 AlaAlaThrProGlnProValLysThrLysGlyValLysArGlyValAspThrThr 295
 ||:|||||
 Db 2512 -----CAGCCCTTGGCCCAAGAAAAAGCGCTTAAGCGAAAGAGATACTAC 2559
 QY 296 ThrProThrThr11e-----AspPro11eHisGluProProSerLeu 309
 ||:|||||
 Db 2560 ACCCTACACCTACAGCATCTTGCTCTGTTCTCCAGCTACCGCTCTGGGAGTCTT 2619
 QY 310 ProProGluProLysThrThrLysLeu--GlyGlnArGluSerSerArGProVal 328
 ||:|||||
 Db 2620 -----GAGCCTAAGGACAGACGGCTTCCCTCATGCTAGAGAGAGTGTGCGCCATC 2673
 QY 329 LysProProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSer 348
 ||:|||||
 Db 2674 AACCCCCACAGCAAGACCTTGCTGACTCTCAGCAACAACACAGAGCTTAAAGAAAGGA 2733
 QY 349 LysValSerGluGlnLeuLysCysCysSerGly11eLeuLysGluMetPheAlaLysLys 368
 ||:|||||
 Db 2734 AACCTTTCAGAACAGTTAAACATTCGCAATGGCATTTTGAAGAGATTACTCTTAAGAG 2793
 QY 369 HisAlaAlaTyrAlaTyrProPheTyrLysProValAspValGluAlaLeuGlyLeuHis 388
 ||:|||||
 Db 2794 CAGTGTGCTCATGCTGTGGCTTTCTATTAACCAAGTATGATCTTCTGCACTTGGCTGAT 2853
 QY 389 AspDtyrCysAsp11eLeuLysHisProMetAspMetSerThr11eLysSerLysLeuGlu 408
 ||:|||||
 Db 2854 GACTACCATATCATATTAGAACCCCATGACTGACCTGCAAGCGGAAAGATGAG 2913
 QY 409 AlaArgGluTyrArgAspAlaGlnGluPheGlyValAspValArgLeuMetPheSerAsn 428
 ||:|||||
 Db 2914 AACCGTATTAACGGGATGACAGAGTTTGGCTGATGATGACGGCTTAATGTTCTCAAC 2973
 QY 429 CysTyrLysTyrAsnProProAspHisGluValValAlaMetAlaArgLysLeuGlnAsp 448
 ||:|||||
 Db 2974 TGTATTAAGTACATATCCCAAGATCAAGATGTTGTGCAATGGACCAAGGACTACAGAT 3033
 QY 449 ValPheGluMetCysPheAlaLysMetProAspGluProGluGluProValValAlaVal 468
 ||:|||||
 Db 3034 GTATTGTGATTCGGTATGCGCAAGATCCAGATGAACCATAGAACAGGAGCCCTTAACA 3093
 QY 469 SerSerProAlaValProProPro--ThrLysValValAlaProProSerSerSerAsp 487

Db 3094 GTCTTACTGCGCAATGCCCCCTGGCTTGCCCAATCGTCTTACAGATCTTCAATGAGGA 3153
 QY 488 SerSerSerSerSerSer-----Asp 495
 ||:|||||
 Db 3154 AGTAGAGTACAGAGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 3213
 QY 496 SerAspSerSerThrAspAspSerGluGluValArgAlaGlnArgLeuAlaGluLeuGln 515
 ||:|||||
 Db 3214 GAAGAGGTGAAGAGCTCAGACTCAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3273
 QY 516 GluGlnLeuLysAlaValHisGluGlnLeuAlaAlaLeuSerGlnProGlnGlnAsnLys 535
 ||:|||||
 Db 3274 GAACAGCTTCGGGACAGTACATGAACAATGCGCTGCTGTGCCAGAGGTCCAAATATCAAG 3333
 QY 536 ProLysLysLysGluLysAspLysLysLysLysLys-----GluLysHis 551
 ||:|||||
 Db 3334 CCCAAGAGGAAA--AGAGAGAAAAGAGAAAAGAGAAACGAGAGAGAGAGAGAGAGAGCAT 3390
 QY 552 LysArgLysGluGluValGluGluAsnLysLysSer--LysAlaLysGluProPro-- 569
 ||:|||||
 Db 3391 CGAGGCGAGCTGGGGCCGATGAAGATGAACAAGGGCTTAAGGCCACCCGCCACCTCA 3450
 QY 570 ProLysLysThrLysLysAsnAsnSerSerAsnSerAsnValSerLysGluProAla 589
 ||:|||||
 Db 3451 CTTAAGAAAGCTCAAGAAACCAAGTGCAGTGGCGGTGGAGAGTCTCTTAAAGCCCTTCT 3510
 QY 590 -----PrometLysSerLys 594
 ||:|||||
 Db 3511 GGCTTTGACCTTTCGAGAGAGTGGACCAACAGCTCCCAAAAAGCCACAAAGAGAGCC 3570
 QY 595 ProPro-----ProThr--TyrGluSerGluGluGluAspLysCysLysPrometSer 611
 ||:|||||
 Db 3571 CCACCTGCCCTGCTCAGATTATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3630
 QY 612 TyrGluGluLysArgLysLysSerLysAsp11eAsnLysLeuProGluGluLysLeuGly 631
 ||:|||||
 Db 3631 TACGATGAGAAAGCGCAGCTGAGCTGACATCAACAAATTACTCTGGGAGAGAGCTGGGC 3690
 QY 632 ArgValValHis11e11eGlnSerArgLysProSerLeuLysAsnSerAsnProAspGlu 651
 ||:|||||
 Db 3691 CGAGTTGTGATATTAATCCAAAGCGAGGAGCCCTCTTTCAGTATTCAAAACCAAGAGAG 3750
 QY 652 11eGlu11eAspPheGluThrLeuLysProSerThrLeuArgGluLeuGlyAlaLeuCys 671
 ||:|||||
 Db 3751 ATTGAATGATTTTAAACATCTCAAGCATCCACATTAGAGAGCTTGA-GCGTATGT 3809
 QY 672 HisLeuLeu 674
 ||:|||||
 Db 3810 CTTTCTCTG 3818
 RESULT 15
 BC032124 2548 bp mRNA linear PRI 26-JUN-2002
 LOCUS BC032124
 DEFINITION Homo sapiens, similar to bromodomain containing 3, clone MGC:29612
 IMAGE:485640, mRNA, complete cds.
 ACCESSION BC032124
 VERSION BC032124.1 GI:21594669
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (baaes 1 to 2548)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 CONTACT Contact: MGC help desk
 COMMENT Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mc@nih.gov

Ahler, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stancirpop, S., Thomas, P.J., Touchman, J.W.,
 Tureseon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
 Series: IRAL Plate: 39 Row: 9 Column: 22
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 12408642.

FEATURES

Source

1. 2548
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:29612 IMAGE:4856840"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH MGC_15"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 308..1978
 /codon_start=1
 /product="Similar to bromodomain containing 3"
 /protein_id="AAH32124.1"
 /db_xref="GI:21594670"
 /translation="MSTATTVAAPGIPATTEGPVNPPEVSNSEKPKRKNQLOLYON
 VVVTTLKHOFPAMPFYOPVDAIKLNPDKHINPDMGIIKRLNNYWSSECM
 QDEFVTLNCTYIYNKPRDILVMAQALEKIFLQVAPQPEVELPPARGKSRKA
 AGASACTOOAAVSVPATFQSVPTVQTVIATPVPTTAVTAVSVPPAAV
 PPPRATPIVPPRPPPVKKYKGRADPTTPTSAITSRSSPPPLSDPKQKV
 ARBSGGRPIPKPKDLDEGVPOHAKGKGLSHLYCOSILREMLSKGAAYAMP
 YKPDVALELDHNDIHKPMDSITVRKMDGEYDQAGFADVLNFSNCKINP
 PDHEVVMARKLDQVFEMRPAKMPDEVEAPALPAPAPAPVSGKESRSSESSDS
 GSSDEERATRLAELOELQALVAHEQALSOAVNPKKKKEKKEKKKDEKEXK
 EKHKVKAEEKKAKVAPAKQAOQKAPAKKANSTTTAGRDHFLTCGV"

CDS

BASE COUNT 641 a 754 c 753 g 400 t
 ORIGIN

Alignment Scores:

Pred. No.: 7,82e-69 Length: 2548
 Score: 1877.50 Matches: 408
 Percent Similarity: 61.33% Conservatave: 63
 Best Local Similarity: 53.12% Mismatches: 152
 Query Match: 51.59% Indels: 146
 Gaps: 18

US-09-700-590a-22 (1-688) x BC032124 (1-2548)

QY 25 GlnMetSerThrThrGlnAlaGlnAlaGlnProGln---ProAlaAsnAlaAlaSerThr 43
 Db 305 AGGATGTCTCCACCGCACAGTCGCCCGCGGGGATCCCGCGACCCCGGCGCTGTG 364
 QY 44 AsnProProProGluThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeu 63
 Db 365 AACCCACCCCGCGAGGCTCTCCAAAGCCCGCGAGCGCGGACCAAGCAAGCTG 424
 QY 64 GlnTyrLeuLeuArgValValLeuLysThrLeuTyrLysHisGlnPheAlaTrpProPhe 83
 Db 425 CAGTACATGACGAATGTGTGTGAAGCGCTCTGGAACAACAAGTTCGCCCTGCTTC 484
 QY 84 GlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleIleLysThr 103

Db 485 TACGACCCCGTGCAGCGCAATCAATTTGAACCTGCCGATTATCATTAATTTAAAC 544
 QY 104 ProMetAspMetGlyThrIleLysLysArgLeuGlnAsnLysTyrTrpAsnAlaGln 123
 Db 545 CCAATGATATGGGAGCTATTAAAGAGAGACTAGAAATAATATTATTATGTGAGTCAAGC 604
 QY 124 GluCyAlaIleGlnAspPheAsnThrMetPheThrAsnCySerTyrIleTyrAsnLysProGly 143
 Db 605 GAATGTATGAGGACTTCAACACCACTGTTTACAAATTTGTTACATTATTAACAAGCCACA 664
 QY 144 AspAspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGln 163
 Db 665 GATACACATAGTGTCTTAAGGCCCAAGCTTTAGAGAAATTTTCTACAAAAAGTGGCCACG 724
 QY 164 LeuProThrGlnGlnThrGlnIleMetIleValGlnAlaLysGlyArgGlyArg---- 181
 Db 725 ATGCCCAAGAGMAAGTTGATTTATTTACCCCTGCTCCAAAGGGCAAGGTCCGAGACCG 784
 QY 182 -----GlyArgLysGlnThrGlyThrAlaLysProGlyValSerThrValProAsnThr 199
 Db 785 GCTCGGGAGAGCCAGAGCGAGTACACAGCAA-----GTGGCGCGCTGTCTGTCTGTC 838
 QY 200 ThrGlnAlaSerThrProProGlnThrGlnThrPro---GlnProAsnProProProVal 218
 Db 839 TCCCCACCG-----ACCCCTTTCAGAGCGTCCGCCACCGCTC 877
 QY 219 GlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProVal 238
 Db 878 TCCGAGAGCGCC-----GTCACTCGCTGCACACCCCTGTA 910
 QY 239 MetThrValValProProGlnProLeuGlnThrProProProValProProGlnProGln 258
 Db 911 CCAACCATCACTCGCAAC-----GTCACTCGCTGCACACCCACGCTCCGCC 961
 QY 259 ProProProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThr 278
 Db 962 CCACCTCT-----CTGCCACACCCACTCGTCCCGTGTCTCTACG 1006
 QY 279 ProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThrProThr 298
 Db 1007 CCG---CCTGTGTCGCAAGAAAAGGGCGTGAAGCGGAACAGACACACACCTCCACG 1063
 QY 299 -----ThrIleAspProIleHisGlnProProSerLeuProProGlnProLys 314
 Db 1064 ACCTCGGCATCATCTGCCAGCGGAGTGTGCGCCCGCGTGG---TCAGACCCCAAG 1120
 QY 315 ThrThrLysLeuGlyLysArgArgLysSer---SerArgProValLysProProLysLys 333
 Db 1121 CAGGCCAAGTGTGTGCGCCCGCGGAGAGTGTGCGCCGCCCATCAAGCCTCCCAAGAG 1180
 QY 334 AspValProAspSerGlnGlnHisProAlaProGlnLysSerSerLysValSerGlnGln 353
 Db 1181 GACCTGAGAGAGCGGCGAGGTGCGCCCGGAGAGTGTGCGCCGCCCATCAAGCCTCCCAAGAG 1240
 QY 354 LeuLysCySerSerGlyIleLeuLysGlnMetPheAlaLysLysHisAlaIleTyrAla 373
 Db 1241 CTGGCTACTGCGACAGATCTCTCAGGAGATGCTATTCAGAAAGACAGCGCGCTACGCC 1300
 QY 374 TrpProPheTyrLysProValAspValGlnAlaLeuGlnLysHisAspTyrCyAspIle 393
 Db 1301 TGGCCCTTCTACAGCCAGTGTGATGCGGAGCGCTTGAGCTGCACAGCTACAGACATC 1360
 QY 394 IleLysHisProMetAspMetSerThrIleLysSerLysLeuGlnAlaArgGlnTyrArg 413
 Db 1361 ATCAAGACCCGAGTGAAGCTAGACACCGTGAAGAAAGAAAGATGAGTGGCCGAGATCAACA 1420
 QY 414 AspAlaGlnGlnPheGlnAlaAspValArgLeuMetPheSerAsnCySerLysTyrAsn 433
 Db 1421 GACGACAGCGGCTTGTGCTGATGTCGCGGTGATGTTCTGAAATTCGTAATAATACAT 1480
 QY 434 ProProAspHisGlnValValAlaMetAlaArgLysLeuGlnAspValPheGlnMetArg 453
 Db 1481 CCCCAGACCAAGAGTGTGCGCATGCGCCGGAAGCTCCAGAGCGTGTGAGATGAGG 1540

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 18:42:55 ; Search time 369 Seconds
(without alignments)
4198.848 Million cell updates/sec

Title: US-09-700-590a-22

Perfect score: 3639
Sequence: 1 MSASGPGTRRLNLPVMDG.....ALCHLLFAEKETFKRLKM 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlD
-O=/Gcn2.1/USPTO-spool/US9700590/runat_14032003_140629_3214/app_query.fasta_1.839
-DB=N_Geneseq_101002 -QFMT=fastap -SUFLTX=eng -MINMATCH=50.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9700590 @Gcn2.1 396 @runat_14032003_140629_3214 -NCPU=6 -ICPU=3
-NO_XLPPX -NO_NMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -IONGLOP -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3639	100.0	2329	AA256719	Human transmembran
2	1888	51.9	4664	ABK83660	Human cDNA differe
3	1888	51.9	4664	ABN95159	Gene #1657 used to
4	1501.5	41.3	3104	ABN68343	Transcriptional re
5	1501.5	41.3	3106	AAK40172	WO9904265 seq id N
6	1450	39.8	1884	AAH18591	Human cDNA sequenc
7	1377	37.8	6776	ABL04337	Drosophila melanog
8	1237.5	34.0	1741	ABK84162	Human cDNA differe
9	1161.5	31.9	1862	ABV25740	Human prostate exp
10	981	27.0	667	AA503029	Human diagnostic a
11	688	18.9	746	AA526193	Human cDNA encodin
12	564	15.5	424	AA526606	Human cDNA encodin
13	484.5	13.3	450	AAZ80432	Human colon cancer
14	457	12.6	560	ABQ60244	Human cDNA encodin
15	430.5	11.8	642	AAK29128	Polynucleotide RTP
16	430	11.8	1542	ABL23049	Drosophila melanog
17	430	11.8	3542	ABL23048	Drosophila melanog
18	411.5	11.3	559	AAA43730	Mouse secreted exp
19	391	10.7	352	AAK66021	Novel human polyu
20	344	9.5	2247	AAK69456	Arabidopsis thalia
21	323.5	8.9	7326	AAV10093	Sequence encoding
22	323.5	8.9	7326	AAK62948	Cellular transcrip
23	323.5	8.9	7326	AAA15554	CREB binding prote
24	323.5	8.9	7326	AAA15554	Mouse nuclear CREB
25	318.5	8.8	807	ABL19177	Drosophila melanog
26	318.5	8.8	2807	ABL19176	Drosophila melanog
27	318.5	8.8	5017	ABL15790	Drosophila melanog
28	308.5	8.5	547	AA557537	cDNA #213 encoding
29	306	8.4	2442	AAK50976	Arabidopsis thalia
30	305	8.4	9046	AAK07792	Transcription fact
31	305	8.4	9046	AAV10092	Human p300 gene SE
32	305	8.4	9046	ABL49723	Human prostate exp
33	303	8.3	597	ABV57572	Human prostate exp
34	302	8.3	209	AAV87340	EST clone BM2. Ho
35	302	8.3	881	AAK56273	Pinus radiata tran
36	300.5	8.3	6743	AAK65172	DNA encoding novel
37	300.5	8.3	6743	AAK88306	Drosophila melanog
38	282	7.7	1445	AAK47126	Arabidopsis thalia
39	282	7.7	3593	ABL09563	Drosophila melanog
40	281.5	7.7	6161	ABL09562	Drosophila melanog
41	281.5	7.7	6161	ABL09562	Arabidopsis thalia
42	277.5	7.6	2014	AAK45013	S. avermitilis ave
43	275.5	7.6	30690	AAA92301	Streptomyces averm
44	275.5	7.6	30690	AAH79277	Drosophila melanog
45	274	7.5	14648	ABL16120	

ALIGNMENTS

RESULT 1
AA256719
ID AA256719 standard; cDNA; 2329 BP.

AC AA256719;
XX
DT 23-MAR-2000 (first entry)

DE Human transmembrane protein HTRMPN-22 encoding cDNA.

KW Human; transmembrane protein; HTRMPN; diagnosis; immunospecific;

KW antiproliferative; neuroprotective; immune disorder;

KW reproductive disorder; smooth muscle disorder; neurological disorder;

KW gastrointestinal disorder; developmental disorder;

XX cell proliferative disorder; ss.

XX Homo sapiens.

XX

PN WO961471-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 28-MAY-1999; 99MO-US11904.
 XX
 XX 29-MAY-1998; 98US-0087260.
 PR 02-JUL-1998; 98US-0091674.
 PR 02-OCT-1998; 98US-0102954.
 PR 24-NOV-1998; 98US-0109869.
 XX
 XX (INCYTE) INCYTE PHARM INC.
 XX
 PI Tang YT, Lai P, Hillman JL, Yue H, Guegler KD, Corley NC;
 PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
 PI An-Young J;
 XX
 DR WPI; 2000-072605/06.
 DR P-PSDB; AAY57898.
 XX
 PT Proteins, polynucleotides, vectors, host cells and antibodies used to
 PT diagnose, treat or prevent immune, reproductive, smooth muscle,
 PT neurological, gastrointestinal, developmental and cell proliferative
 PT disorders -
 XX
 XX Claim 9; Page 191; 229pp; English.
 XX
 CC AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human
 CC transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively.
 CC The transmembrane protein have immunospecific, antiproliferative and
 CC neuroprotective activities. The human transmembrane proteins,
 CC polynucleotides encoding them and other compositions and methods from
 CC the present invention, can be used for the diagnosis, treatment or
 CC prevention of immune, reproductive, smooth muscle, neurological,
 CC gastrointestinal, developmental and cell proliferative disorders. The
 CC HTMPN's can be used to treat or prevent disorders associated with a
 CC decreased expression or activity of HTMPN.
 XX
 SO Sequence 2329 BP; 674 A; 691 C; 569 G; 382 T; 13 other;
 Alignment Scores:
 Pred. No: 1.09e-166 Length: 2329
 Score: 3639.00 Matches: 688
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-700-590a-22 (1-688) x AAZ56719 (1-2329)
 QY 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 DB 16 ATGCTGGGAGAGGAGGCGCTGGGACGAGATTGAGAAATCTGCCGATATGGGGATGGA 75
 QY 21 LeuGluThrSerGlnMetSerThrThrglnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 DB 76 CTAGAACTTCCCAATGCTTCAACACACAGGCGGCGCAACCCACGACGCAACCA 135
 QY 41 AlaSerThrAsnProProProProGluThrSerAsnProAsnLysProLysArgGlnThr 60
 DB 136 GCCAGCAACCAACCCCGCCCGCCAGACCTTCAACCTTCAACCCCGCAAGGAGGAGCC 195
 QY 61 AsnGlnLeuGlnThrLeuLeuArgValValLeuLysThrLeuTyrLysHisGlnPheAla 80
 DB 196 AACCAACTGCAATACCTCTCAGAGTGGCTCAAGACCTATGGAACACCAAGTTTGA 255
 QY 81 TyrProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 DB 256 TGGCTTTCCAGCAGCCTCTGATGCCGTCAAGCTGAACTCTCCGATTAATAGATC 315
 QY 101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGluAsnAsnTyrTyrTrp 120
 DB 316 ATTAAGCGCTATGATATGGAACTAAAGAGCGCTTGGAAAAACAATATTACTGG 375

QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnGlyTyrIleTyrAsn 140
 DB 376 AATGCTCAGGAAAGTATCCAGGACCTTCAACACTATGTTTCAAAATTGTACTACTACAC 435
 QY 141 LysProGlyAspAspIleValLeuMetAlaGlnAlaLeuGluLysLeuPheLeuGlnLys 160
 DB 436 AAGCTGTGAGATGACATAGTCTTAAATGGCAGAACTCTTGAAAAAGCTCTTTCGCAAAA 495
 QY 161 IleAsnGluLeuProThrThrGluGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 DB 496 ATAAATAGCTTACCCACAGAAAGAAACCGAGATCATGATATGTCAGGCAAAAGAGAGGA 555
 QY 181 ArgGlyArgLysGluThrThrAlaLysProGlyValSerThrValProAsnThrThr 200
 DB 556 CGTGGAGGAAAGAAACAGGAGACAGCAAACTGGCGTTTCCACGGTACCAACACAACT 615
 QY 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsnProProProValGlnAla 220
 DB 616 CAAGCATCGACTCTCCCGCAGACCCACAGACCCCTCAGCCGAATCTCTCTGTGACAGGC 675
 QY 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMetThr 240
 DB 676 AGCCTCACCCCTTCCCTGCGTCACCCCGAGCTCATCTGTCAGACCCCTGTATACA 735
 QY 241 ValValProProGlnProLeuGlnThrProProValProProGlnProGlnProPro 260
 DB 736 GTGTGCTCTCCCGCAGCAGTACAGACGCGCCGCGAGTGGCCCGCCAGCAACCCCA 795
 QY 261 ProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaIleThrProGln 280
 DB 796 CCCCCTCAGTCCCGCCGCGGTACAGAGCACCACCCATCATCTGGGGCACCACCCAG 855
 QY 281 ProValThrThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle 300
 DB 856 CCTGTGAAGACAAAGAGAGAGTGAAGAGAAAGCAACACACACCCACCACTT 915
 QY 301 AspProIleHisGluProProSerLeuProProGluProLysThrThrLysLeuGlyGln 320
 DB 916 GACCCATTCACGAGCACCCTCGCTGCCCGGAGCCCAAGACCAACAGCTGGGCGAG 975
 QY 321 ArgArgLysSerSerArgProValLysProProLysLysAspValProAspSerGlnGln 340
 DB 976 CGGCGGAGAGACACCGCGCTGTGAAACCTTCAAAAGAGACGCGCCGCTCTCACAG 1035
 QY 341 HisProAlaProGluLysSerSerLysValSerGluGlnLeuLysCysCysSerGlyIle 360
 DB 1036 CACCCAGACCAAGAGAGAGCAGCAGGCTCTGGAGCAGCTCAAGTCTGACGGGCAATC 1095
 QY 361 LeuLysGluMetPheAlaLysLysHisAlaAlaTyrAlaTrpProPheTyrLysProVal 380
 DB 1096 CTCAAGGAGATGTTTGGCAAGAAAGCACCGCGCTAGCCCTTGCCCTTCAACACCTGTG 1155
 QY 381 AspValGluAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisProMetAspMet 400
 DB 1156 GACGTGAGGACACTGGGCGCTACACGACTACTGTGACTCATCAAGCAACCCCATGAGATG 1215
 QY 401 SerThrIleLysSerLysLeuGluAlaArgGluTyrArgAspAlaGlnGluPheGlyVala 420
 DB 1216 AGCAAAATCAAGTCTTAACTGAGAGGCCGTGAGTACCTGATGCTCGAGAGTTGGTGTCT 1275
 QY 421 AspValArgLeuMetPheSerAsnGlyTyrLysTyrAsnProProAspHisGluValVala 440
 DB 1276 GACGTCGATTTGATGTTCTCCAACTGCTATATAGTACAACCTCTGACATGAGAGTGTG 1335
 QY 441 AlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMetProAspGlu 460
 DB 1336 GCCATGGCGCCGCAAGCTCCAGAGATGTGTGAAATGGCTTTGGCAAGATGCCGAGAG 1395
 QY 461 ProGluGluProValValAlaValSerSerProAlaValProProProThrLysValVal 480
 DB 1396 CCGAGGAGCGAGTGTGGCGGTCTCTCCCGGCAAGTGGCCCTCCACCAAGAGTTGTG 1455

[illegible]

PA	(GENE-) GENE LOGIC INC.
Pt	
Pt	Detecting granulocyte activation by detecting differential expression
Pt	of genes associated with granulocyte activation, which serves as
Pt	diagnostic markers that is useful for monitoring disease states and
Pt	drug toxicity -
XX	
PS	Claim 1; SEQ ID No 231; 114pp; English.
XX	
CC	The invention relates to detecting (M1) granulocyte (GC) activation
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC	DNA chip analysis as given in the specification, and comparing
CC	the expression level to an expression level in an unactivated
CC	Gc, where differential expression of Gs is indicative of GCA.
CC	Also included are modulating (M2) GA by contacting GC with an agent
CC	that alters the expression of at least one gene in Gs; (2) screening (M3)
CC	for an agent capable of modulating GCA or an inflammation (especially
CC	chronic) in a tissue, an allergic response in a subject, exposure of a
CC	subject to a pathogen or sterile inflammatory disease using the
CC	gene expression profile; (3) detecting (M4) an inflammation (especially
CC	chronic) in a tissue, an allergic response in a subject, exposure of a
CC	subject to a pathogen or sterile inflammatory disease, by detecting the
CC	level of expression in a sample of the tissue of gene(s) from Gs, where
CC	the level of expression of the gene is indicative of inflammation;
CC	(4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC	an allergic response in a subject, exposure of a subject to a pathogen
CC	or sterile inflammatory disease, by contacting a tissue having
CC	inflammation with an agent that modulates the expression of gene(s)
CC	from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC	modulating GA; M3 is useful for screening an agent capable of modulating
CC	GCA preferably in an inflammation in a tissue; M4 is useful for
CC	detecting an inflammation (especially chronic) in a tissue, an allergic
CC	response in a subject, exposure of a subject to a pathogen or sterile
CC	inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC	glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC	reflexion injury, ARDS, adult respiratory distress syndrome,
CC	infectious bowel disease, Crohn's disease, ulcerative colitis,
CC	periodontal disease; also bacterial infection, viral infection,
CC	parasitic infection, protozoal infection, fungal infection and M5 is
CC	useful for treating one of the above conditions. The present
CC	sequence represents a gene differentially expressed in granulocytes.
CC	Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic
CC	format directly from WIPo at
CC	ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 4664 BP; 1084 A; 1336 C; 1293 G; 951 T; 0 other;
SS	
Alignment Scores:	
Pred. No.:	3,34e-82 Length: 4664
Score:	1888.00 Matches: 404
Percent Similarity:	67.08% Conservative: 81
Best Local Similarity:	55.88% Mismatches: 142
Query Match:	51.88% Indels: 97
DB:	24 Gaps: 20
US-09-700-590A-22 (1-688) x ABK83660 (1-4664)	
Oy	5 serglyProglYThrArgLeuArgAsnLeuProValMetGlYAspGlYleuGlunThrSer 24
Db	::: :::: : : ::
1780 GCACGACCAGGAAAGATGCCAAAACCTCCTCTTGTATAGAGGGCTTTAGAGGCC 1839	
Oy	25 GlmMetserThrTrgInAlaGlnAlaGlnProGlnProAlaAsnAlaIaslerThrsn 44
Db	:::: : : ::
1840 ACAATGGCTTCGGGCTCGT---TTGCAACTTACCCCTGCC-----AAC 1881	
Oy	45 ProProProProGluThrSerAsnProAlaenLySProlYArGrInThrAsnGlnenun 64
Db	:::: : : ::
1882 CCACGACCCCCGGAGGTGTCTCAATCCCAAAGAAGCCGAGCGATTACCAACCGCTGGAA 1941	

Oy	65	TyLeuLeuAargValValLeuLysThrLeuTrpLysHisGlnPheLeuAlaTrpProGln	84
Db	1942	TACCTACACAAAGGTAAGTAAGTAAGAGCCCTCTGGAAACATCAAGTTCGACATGGCCATTCGCG	2001
Oy	85	GlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleLysThrPro	104
Db	2002	CAGCTGTGGATGTGTCCTAAACTGGCTACCGGATATATCAGAAATTATTAACAAGGCT	2061
Oy	105	MetAspMetGlyThrIleLysLysArgLeuGlnAsnAsnTyrTyrTrpAsnAlaGlnGlu	124
Db	2062	ATGGACATGGGTACTATTATTAAGAGAGACCTTGAAAACAAATTATTAATGGCTCTTCACAG	2121
Oy	125	CysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsnLysProGlyAsp	144
Db	2122	TGTATGCAAGATTTTAAATACATGTTTACCAACTGTTCATTTACCAACAAGCCACTGAT	2181
Oy	145	AspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGluLeu	164
Db	2182	GATATTGTCTTAATGGCAACAAGCTGGAAAAATATTCTTACAGAAAGTTGCATCAATG	2241
Oy	165	ProThrGlnGlnThrGlnLysIleMetLysIleValGlnAlaLysGly-----ArgGly---	180
Db	2242	CCACAGAAAGAAACAAGAGCTGTGTAGTACATCCCTTAAGAACACCAAGAAGGGGGCC	2301
Oy	181	-----ArgGlyArgLysGlnThrGlyThrAlaLysProGlyValSerThr	195
Db	2302	AAGTTGGACGGCTCCACGGGAGATGTTTACAAGTGGCCCATAGGTGCTCCCTCTCT	2361
Oy	196	ValProAsnThrThrGlnAlaSerThrProGlnIleThrGlnThrProGlnProAsnPro	215
Db	2362	GTGTCAACACAGGCCGTGTATCTCTCCACCTGAGATA-----CCT	2403
Oy	216	ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln	235
Db	2404	ACCACTGCTCCAAACATTTCCCAACCA-----TCAGTCAATTCC	2442
Oy	236	ThrProValMetThrValValProProGlnProLeuGlnThrProProProValProPro	255
Db	2443	TCTCCACTTCCAAAGCTCCGGAC-----TCTGGTGAACCCCGGCTCTTGGCT	2490
Oy	256	GlnProGlnProProProAlaProAlaProGlnProValGlnSerHisProProIleIle	275
Db	2491	GTTACTGCAGCTCTCTCCAGC-----	2511
Oy	276	AlaAlaThrProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThr	295
Db	2512	-----CAGCCCTTGGCCCAAAABAAAAGGGGTAAAGCGAAACAAATACTAC	2559
Oy	296	ThrProThrThrIle-----AspProIleHisGlnProProSerLeu	309
Db	2560	ACCCTTACACCTTACAGGCATCTTGCTGCTCTTCAGTCTTCAGCTGAGCCCTCTGGAGTCTT	2619
Oy	310	ProProGlnProLysThrThrLysLeu---GlyGlnArgArgLysSerArgProVal	328
Db	2620	-----GGGCTTAAAGCAGACAGGCTTCCCTATATGGTATGAAGAGTGTGCGCCATC	2673
Oy	329	LysProProLysLysAspValProAspSerGlnGlnHisIleProAlaProGlnLysSerSer	348
Db	2674	AAGCCCCACCCAAAGACTTGGCTCTCAGCAACAAACACAGAGCTCTTAAGAAAGA	2733
Oy	349	LysValSerGlnGlnLeuLysCysCysSerGlyTyrLeuLysGlnMetPheAlaLysLys	368
Db	2734	AAGCTTACAGAAAGTTAAACATTTGCAAGGACATTTTGAAGAGATTACTCTTAAGAG	2793
Oy	369	HisAlaAlaTyrAlaTrpProPheTyrLysProValAspValGlnAlaLeuGlnLysHis	388
Db	2794	CATGCTGCTTATGCTTGGCTTTTATTAACACAGTGAATCTTCTCAGTTGGCTGAT	2853
Oy	389	AspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGln	408
Db	2854	GACTCACTGATGATCATTTAAGACCCCATGAGCTCTGACACTGTCAAGCGAAGATGAG	2913

OY	409	AlaArgIuIuYrArygAspIaIagInguIuPheGlyAlaaspValArIeUmetPhSerSer	428
OY	409	AlaArgIuIuYrArygAspIaIagInguIuPheGlyAlaaspValArIeUmetPhSerSer	428
Db	2914	AACCGTAATTCACGGGATGACAGGAGTTTGCTGCTGATGATGAGCTTATGTTCTCCAA	2973
OY	429	CysTyrIysTyrAsnProProAspHisGluValAlaMetAlaArgIuLeuGlnAsp	448
Db	2974	TGCTATAAGTACATCCCCAGATACAGATGTTGGCCATGGCAGAAAGCTACAGAT	3033
OY	449	ValPheGluMetArgPheAlaIysMetProAspGluProGluGluProValAlaAla	468
Db	3034	GTATTAGTATCCGTTATGCGAAGATGCAAGATGAAACCTAGAAACGAGGCTTTACA	3093
OY	469	SerSerProAlaValProProPro---ThrIysValAlaIaProProSerSerSerAsp	487
Db	3094	GTCCTACTCGCATGCCCTCGCTGGCTTGCCCAATGCTTTCAGAGTCTCCAGTAGGAA	3153
OY	488	SerSerSerAspSerSerSer-----Asp	495
Db	3154	AGTAGCAGCTAGAGACTCTCTTGAGAAAGAGAGAGAGAAAGATGAGAGAGACGAGAGAA	3213
OY	496	SerAspSerSerThrAspAspSerGluGluGluArgAlaGlnArgLeuAlaGluLeuGln	515
Db	3214	GAAAGAGTGAAAGCTCAGACTCAGAGAGAAAGGGCTCATCGCTTGCGAAGACTTACAG	3273
OY	516	GluGluIuIuYsAlaValHisGluGluIuMetAlaAlaIeuSerGluProGluGlnAsnIys	535
Db	3274	GAAACGCTTCGGGAGACAGTACAGTAAACAACGCTGCTCTGCCAGGCTCCATATTCAG	3333
OY	536	ProIysIysIysGluIuYsAspIuIysGluIuIysIysIys-----GluIysHis	551
Db	3334	CCCAAGAGGAAA---ACAGGAAAAAAGAAAAAGAAAGAAACGAGAGAGAGACAT	3390
OY	552	LysArgIysGluGluIuValGluGluIuMetIysIysSer--LysAlaIalysGluProPro--	569
Db	3391	CGAGGCGCAGAGCTGGGGCCCATGAAGATGACAAAGGGGCTTAGGGGCAACCCGCCACTCAA	3450
OY	570	ProIysIysThrIuIysIysAsnAsnSerSerAsnSerAsnValSerIysIysGluProAla	589
Db	3451	CCTAAGAACTCCAAAGAAAGCAAGTGCGAGTGGGGGTGCAGTGCTGCTTAAAGCCCTTCT	3510
OY	590	-----ProMetIysSerIys	594
Db	3511	GGCTTTGGACCTTCTGGAGAAAGTGGCACCAAGCTCCCAAAAAGGCCACAAGACACAGC	3570
OY	595	ProPro-----ProThr--TyrgIuSerGluGluIuAspIysCysIysProMetSer	611
Db	3571	CCACCTGCGCTGACCTAACAGGTATGATTACAGAGAGAGAGAAAGACAGAGCCCATAGT	3630
OY	612	TyrGluGluIuYsArgIuIuSerIeuSerIeuAspIleAsnIysIeuProGluIuIuIysIeuGly	631
Db	3631	TACATGTAGAAAGCGGACGCTGAGCCCTGGACATCAACAAATTCTCGGGAGAGAACTGGGC	3690
OY	632	ArgValAlaHisIleIleGlnIserArgIuProSerIeuIysAsnSerAsnProAspGlu	651
Db	3691	CGAGTGTGCATATATATCCCAAGCCAGAGAGCCCTCTTTACGGATTCAAACCCAGAAAG	3750
OY	652	IleGluIleAspPheGluThrIuIuIuIysProSerThrIeuArgIuIuIuGlyAlaIeuCys	671
Db	3751	ATTGAGATTGATTTTGAACAACCTCAAGCCATCCACTTAGAGAGACTTGA-GCGCTATGT	3809
OY	672	HisIeuIeu 674	
Db	3810	CCTTTCCG 3818	
RESULT 3			
ABN95159			
ID	ABN95159 standard; DNA; 4664 BP.		
AC	ABN95159;		
XX			
DT	13-AUG-2002 (first entry)		
XX			
DE	Gene #1657 used to diagnose liver cancer.		

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumor; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 PN WO200229103-A2.
 PD 11-APR-2002.
 PF 02-OCT-2001; 2001WO-US30589.
 PR 02-OCT-2000; 2000US-237054P.
 PA (GENE-) GENE LOGIC INC.
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 DR WPI; 2002-426119/45.
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample.
 PS Claim 1: SEQ ID NO 1657; 298bp; English.
 XX The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytosolic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_poc_sequences.
 SO Sequence 4664 BP; 1084 A; 1336 C; 1293 G; 951 T; 0 other;

Alignment Scores:
 Pred. No.: 3,34e-82 Length: 4664
 Score: 1888.00 Matches: 404
 Percent Similarity: 67.08% Conservative: 81
 Best Local Similarity: 55.88% Mismatches: 142
 Query Match: 51.88% Indels: 97
 Gaps: 20

US-09-700-590A-22 (1-688) x ABN95159 (1-4664)

QY 5 SerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGluThrSer 24
 DB 1780 GCGACGACGAGGAAAGAGATTGCAAAACCTCTCTGTATGAGGGCTTTGAGAGCC 1839
 QY 25 GlMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaIleSerThrAsn 44
 DB 1840 ACAATGGCTTGGTGGCTGCT---TTGCAACTTACCCCTGCC-----AAC 1881
 QY 45 ProProProGluThrThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGln 64
 DB 1882 CCAACGACCCCGGAGGTGTCATCCCAAAAGCAGAGAGATTACCAACGAGCTGCA 1941
 QY 65 TyrLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAlaTrpProPheGln 84
 DB 1942 TACCTACACAGAGTAGTAGAAGGCTCTGTGAAACATCATGTCGATGCCATTCGCG 2001
 QY 85 GlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleIleLysThrPro 104

DB 2002 CAGCCCTGTGAATGCTGTAACCTGGCTTACCGGATTATCACAAAATTTAAACAGCCT 2061
 QY 105 MetAspMetGlyThrIleLysLysArgLeuGluAsnAsnTyrTyrAsnAlaGlnGlu 124
 DB 2062 ATGACATGGGACTACTTTAAGAGAGACTTGAACAACTATTATTATGGCTGCTCAGAG 2121
 QY 125 CysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsnLysProGlyAsp 144
 DB 2122 TGTATGCAAGATTTTAAATACATGTTTACCACTGTTATCATTTTACAAAGCCACTGAT 2181
 QY 145 AspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLysIleAsnGluLeu 164
 DB 2182 GATATTGCTCTAATGACACAAACGCTGGAAGAAAGATTCTTACAGAAAGGTTCATCAAG 2241
 QY 165 ProThrGluGluThrGluIleMetIleValGlnAlaLysGly-----ArgGly--- 180
 DB 2242 CCAAGAAAGAACAGAGAGCTGTAGTACCATTCCTTAAGAAACAGCCACAAAGAGGGGCG 2301
 QY 181 -----ArgLysArgLysGluThrGlyThrAlaLysProGlyLysThr 195
 DB 2302 AAGTTGGACGCGCTCCAGGAGGTGTACCAAGCCCATCAGGTGCTCGCTCTTCT 2361
 QY 196 ValProAsnThrThrGlnAlaSerThrProProGlnThrGlnThrProGlnProAsnPro 215
 DB 2362 GTGTACACACAGCCCTGTATCTCTCCACCTGAGATA-----CCT 2403
 QY 216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235
 DB 2404 ACCACTGTCTCTCAACATTCGCCACCA-----TCAGCATTTTCC 2442
 QY 236 ThrProValMetThrValAlaProProGlnProLeuGlnThrProProProValProPro 255
 DB 2443 TCTCCACTTCTCAAGCTTGAC-----TCGTGGAGACCCCGCTCTGTCT 2490
 QY 256 GlnProGlnProProProAlaProAlaProGlnProValGlnSerHisProProIleIle 275
 DB 2491 GTTACTGACAGCTCTCCAGCC----- 2511
 QY 276 AlaAlaThrProGlnProValLysThrLysGlyValLysArgLysAlaAspThrThr 295
 DB 2512 -----CAGCCCTGTGCAAGAAAGAGGGTAAAGCGGAACAGATCTAC 2559
 QY 296 ThrProThrThrIle-----AspProIleHisGluProProSerLeu 309
 DB 2560 ACCCTTACACCTACAGCCATCTTGCTCTGCTGCTTCCAGCTACGCCCTCTGGAGATCTT 2619
 QY 310 ProProGluProLysThrThrLysLeu---GlyGlnArgArgGluSerSerArgProVal 328
 DB 2620 -----GAGCTTAAGGACAGACAGGCTTCCCTATGCGTATGAGAGAGTGTGCCCCATC 2673
 QY 329 LysProProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSer 348
 DB 2674 AAGCCCCCAGCAGCAAGATCTGCTGACTCTCAGCAACACACAGAGCTCTAAGAAAGGA 2733
 QY 349 LysValSerGluGlnLeuLysCysCysSerGlyIleLeuLysGluMetPheAlaLysLys 368
 DB 2734 AAGCTTTCAGAACAGTTAAACCTTGCAATGCGATTTTGAAGAGTTACTCTTAAGAG 2793
 QY 369 HisAlaAlaTyrAlaTrpProPheTyrLysProValAspValGluAlaLeuGluLysHis 388
 DB 2794 CATGCGCTCATGCTTGCTGCTTCTATAAACAGTGAGTCTCTGCACTGGCTGCAT 2853
 QY 389 AspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGlu 408
 DB 2854 GACTTACCATGACATCTTAAAGCAACCCCATGACCTTGAAGAGTCTTAAAGAGATGAG 2913
 QY 409 AlaArgGluTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsn 428
 DB 2914 AACCGTATTAACGGAGATGACAGAGTTTCTGCTGATGATGAGGCTTATGTTCTCAAC 2973
 QY 429 CysTyrLysTyrAsnProProAspHisGluValAlaMetAlaArgLysLeuGlnAsp 448

Db 2974 TGCTAATAGTAAATCCCCAGATCAGATCTTGTGGCAATGGCAAGAACTACAGAT 3033
 Qy ValPheGluMetArgPheAlaLysMetProAspGluProGluGluProValAlaVal 468
 Db 3034 GTATTGAGTTCCCTTATGCCAAGATGACATGACACCAAGCAAGGCGCTTACCA 3093
 Qy SerSerProAlaValProProPro---ThrLysValValAlaProProSerSerSer 487
 Db 3094 GTCTCTACTGCGCATGCCCGCTGGCTGGCCAAATCGTCTTCAGAGTCTCCAGTGAGAA 3153
 Qy SerSerSerSerSerSerSer-----Asp 495
 Db 3154 AGTAGCGGTGAGAGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 3213
 Qy SerAspSerSerThrAspAspSerGluGluGluGluArgAlaGlnArgLysAlaGluLeuGln 515
 Db 3214 GAAGAGAGTGAAGAGCTGACACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3273
 Qy GluGluLeuLysAlaValAlaGluGluLeuAlaLeuSerGlnProGluGlnAsnLys 535
 Db 3274 GAAACAGCTTCGGGAGTACATGAACAACAGCTGCTGTCTCCAGAGGTCCAAATATCCAA 3333
 Qy ProLysLysLysGluLysAspLysLysLysLysLys-----GluLysHis 551
 Db 3334 CCCAAGAGAGAAA---AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3390
 Qy LysArgLysGluGluValAlaGluLysLysSer---LysAlaLysGluProPro--- 569
 Db 3391 CGAGGCCGAGTGGGGCCGATGAGATGACAAAGGGGCTAGGGGACCCCGCCACCTCAA 3450
 Qy ProLysLysThrLysLysAsnAsnSerSerSerSerValSerLysLysGluProAla 589
 Db 3451 CCTAAGAAGTCCAGAAAGCAAGTGGCAGTGGGGGTGGCAGTGTCTTAGGCCCTTCT 3510
 Qy 590 -----PromLysSerLys 594
 Db 3511 GGCCTTTGACCTTGGAGAGAGTGGCAACAAGCTCCCAAAAAGGCCAACAAAGACGCC 3570
 Qy ProPro-----ProThr---TyrGluSerGluGluGluAspLysCysLysProMetSer 611
 Db 3571 CCACCTGCGCCCTACAGCTTATGATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3630
 Qy TyrGluGluLysArgGluLeuSerLeuAspLysLysLysLysLysLysLysLysLys 631
 Db 3631 TACAGATGAGAGCGGACGCTGAGCTGACATCAACAATTACCTGGGGAGAGAGCTGGGC 3690
 Qy 632 ArgValValHisLeuLeuSerArgLysProSerLeuLysAsnSerSerProAspLys 651
 Db 3691 CGAGTTGTGCATATAATCAAGCCAGGAGCCCTCTTACGATTCAAACCCAGAGAGAG 3750
 Qy IleGluIleAspPheGluThrLeuLysProSerThrLeuArgGluLeuGlyAlaLeuCys 671
 Db 3751 ATTGAGATGATTGTTGAACAACCTCAAGCCATCCACACTTAGAGAGCTTGA-GCGCTATGT 3809
 Qy 672 HisLeuLeu 674
 Db 3810 CCTTCTCTG 3818
 RESULT 4
 AAV68343
 ID AAV68343 standard; cDNA to mRNA; 3104 BP.
 XX AC AAV68343;
 XX 05-MAR-1999 (first entry)
 XX Transcriptional regulatory factor RING3 encoding cDNA.
 KW Human; transcriptional regulatory factor; RING3; TSB; cancer;
 KW testis specific bromodomain; testicular cell proliferation; ds.
 XX Homo sapiens.
 OS XX

FH Key Location/Qualifiers
 FT CDS 106..2949
 TT /*tag= a
 PN MO9848015-A1.
 XX 29-OCT-1998.
 PD 17-APR-1998; 98MO-JD01782.
 XX 18-APR-1997; 97JP-0116402.
 PR (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA Jones MH;
 XX WPI; 1998-583658/49.
 DR P-PDB; AAW81168.
 XX Transcriptional regulator gene containing bromodomain sequence - may
 PT be expressed in testis tissue and is useful in treatment of cancer
 PT and other proliferative disorders
 XX Claim 2; Page 24-31; 42pp; Japanese.
 CC The present sequence encodes the human transcriptional regulatory
 CC factor RING3, which is isolated from testicular cells. RING3 contains
 CC a testis specific bromodomain (TSB) which is expressed specifically
 CC in testis tissue and also expressed in certain tumour lines. The
 CC transgenic cells may be used to express RING3 which is a TSB expression
 CC protein. The TSB expression product can be used in the treatment of
 CC cancer and other proliferative disorders, and in screening of compounds
 CC for ability to bind to it (e.g. for use as drugs by modulation of
 CC transcriptional regulation). DNA capable of hybridising to RING3
 CC polynucleotides may be used for construction of probes and primers.
 XX
 SQ Sequence 3104 BP; 1179 A; 560 C; 593 G; 772 T; 0 other;
 Alignment Scores:
 Pred. No.: 8..8e-64 Length: 3104
 Score: 1501.50 Matches: 326
 Percent Similarity: 61.67% Conservative: 81
 Best Local Similarity: 49.39% Mismatches: 160
 Query Match: 41.26% Indels: 94
 DB: Gaps: 13
 US-09-700-590A-22 (1-688) x AAV68343 (1-3104)
 Qy 35 ProGluProAlaAsnAlaAlaSerThrAsnProProProGluThrSerAsnProAsn 54
 Db 115 CCAAGTGCACAAACAGCTATTATTGTTTAACTCTCTCCACCAAGATATATATACTAAG 174
 Qy LysProLysArgGluThrAsnGluLeuGlnTyrLeuLeuArgValAlaLeuLysThrLeu 74
 Db 175 AAAAATGGCGCATTGACAAATCACTTCAGTATCTACAAAAGTTGCTTAAAGATTTA 234
 Qy TyrLysHisGluPheAlaTrpProPheGluGluProValAspAlaValLysLeuAsnLeu 94
 Db 235 TGAAGCATAGATTTTTCATGGCCCTTCAACGTCCTGATGCTGGAACCTAAAGTTG 294
 Qy ProAspTyrTyrLysLysLysLysLysThrProMetAspMetGlyThrLysLysArgLeu 114
 Db 295 CCGATATTATATACCATTTATTAACCAACCAATGATTTAATAATTAAGAACCGCTTG 354
 Qy GluAsnAsnTyrTyrTyrPheAsnAlaGluGluCysLysGluAspPheAsnThrMetPheThr 134
 Db 355 GAGATATAATTTATATGCGAAGGCTTCAGAAATGATATAGAGCTTCATATACATTTCTCA 414
 Qy AsnCysTyrLysLysLysLysLysProGlyAspAspLysValLeuMetAlaGluAlaLeuGlu 154
 Db 415 AATGTTATTATATATACCAAGCTCGAGATCACAATGTTCTTAAGGCAACAAGCTTAGAG 474
 Qy LysLeuPheLeuGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 174


```

Db 475 AACCTGTTATGAGAAATTAATCTCAGATGCCACAAAGAAAGCAAGTTGGGTGTT-- 531
Qy 175 GlnAlaLeuSgLYArgGlyArgGlyLeuThrGlyThrAlaLysProGlyValSer 194
Db 532 -----AAGAAAGAAATCAAGAAAGCACT----- 555
Qy 195 ThrValProAsnThrThrGlnAlaSerThrProGlnThrGlnThrProGlnProAsn 214
Db 556 -----CAACAGAAAT 564
Qy 215 ProProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuLeval 234
Db 565 ATAGCTGTTTCTTCTGTAAAGAAATATCATCCCGCAAGCAAGAAAGATTTTAAAG 624
Qy 235 GlnThrProValMetThrValAlaProGlnProGlnProLeuGlnThrProProValPro 254
Db 625 CACCAAGAAATTCCTTCTGTATTCTTAAGACATCTATT----- 663
Qy 255 ProGlnProGlnProProAlaProAlaProGlnProValGlnSerHisProProIle 274
Db 664 -----TCTCCCTTGAACGTGTACGAGGAGCTTCAGTC 696
Qy 275 IleAlaAlaThrProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThr 294
Db 697 AACCTCAGATTCACAAATCGCGCCCAAGTTACAAAGAGTGAAGAAAGCAAGATACA 756
Qy 295 ThrThrProThrThrIleAspProIleHisGlnProProSerLeuProProGlnProLys 314
Db 757 ACAACCTCTGCAACT--TCAGCAGTTAAAGCAAGTAGTAATTTCTCCA----- 804
Qy 315 ThrThrLysLeuGlyGlnArgArgGlySerSerArgProValLysPro-----ProLys 332
Db 805 -----ACATTCACAGAAATATCAGTGGCACTGCCACTTAAAGAAATATAGCCCAAG 858
Qy 333 LysAspValProAspSerGlnGlnHisProAlaProGlnLysSerSerLysValSerGln 352
Db 859 AATGTTTTCAGATTCCTCAGCAACATATATATGTTGTGAGACTGTTAAAGTAACGAA 918
Qy 353 GlnLeuLysCysCysSerGlyIleLeuLysGlnMetPheAlaLysLysHisAlaAlaTyr 372
Db 919 CAATTAAAGCATGTAGTGAATTTCTTAAGAAATGCTTCAAGAAACATTTTCTCATAT 978
Qy 373 AlaTyrProPheTyrLysProValAspValGlnAlaLeuGlyLeuHisAspTyrCysAsp 392
Db 979 GCATGGCCCTTTATATATCTGTTGACGTTAAATGCTTGGGACTCCATATCTATGAC 1038
Qy 393 IleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGlnAlaArgGluTyr 412
Db 1039 GTTGTCAAAAATCCGATGATCTTGAACATATTAAAGAAATGATTAACAAGATAT 1098
Qy 413 ArgAspAlaGlnLeuPheGlyAlaAspValArgLeuMetPheSerPheCysTyrLysTyr 432
Db 1099 AAGGATGCAATCACTATTGGCGAGATGTAGATTGTTCAATTCCTCAAGATAC 1158
Qy 433 AsnProProAspHisGlnValAlaIleMetAlaArgLysLeuGlnAspValPheGlnMet 452
Db 1159 AATCTCCACATACAGAAATGTGACATATGCAAGAAATGCTTCAAGATGTTTTCGAACG 1218
Qy 453 ArgPheAlaLysMetProAspGluProGlnGlu-----ProValAlaAlaLysSer 470
Db 1219 CATTTCACAAAGATCCGATGAACTGTGAGAGTATGCTTTATGTTACATCAAAACA 1278
Qy 471 ProAlaValProProProThrLysValAlaAlaProProSerSerAspSerSer 490
Db 1279 GAT-----ATCACAGAAACCACTGTAGAGAAACACATAT 1314
Qy 491 AspSerSerSerAspSerAspSerSerThrAspAspSerGlnGlnGlnAlaGlnArg 510
Db 1315 GAAGCTCTCTCTGAAGGAACCTTCT--GATGATTTCTAAGATGAGCAAGTTAAAGCT 1371
Qy 511 LeuAlaGlnLeuGlnGlnGlnLeuLysAlaValHisGlnGlnLeuAlaLeuSerGln 530

```

```

Db 1372 CTTCGAAAGCTTCAGAGCAGCTTAAAGCTGTACATCAACAGCTCCAGGTTTGTCCAA 1431
Qy 531 ---ProGlnGlnAsnLysProLysLysGlyLysAspLysLysGlnLysLysGln 549
Db 1432 GTACCTTTCGTTAGCTAATATAAAGAAAGAAAGACTCTAAAGAAAGAAAGAAAGAA 1491
Qy 550 Lys-----HisLysArgLysGlnGlnValGlnGlnLysLysSer 563
Db 1492 AAGGTTAATTAACAGCAATGAAATCCAGAAATAATGTGTGACCAATGAGGCTTAAAGAA 1551
Qy 564 LysAlaLysGlnProProProLysLysThrLysLysAsnAsnSerSerAsnVal 583
Db 1552 AAGTCCAAAGAAATATGACCAAGAAAGAAAGAAAGCAAGCTTCAATTCGT----- 1599
Qy 584 SerLysLysGlnProAlaProMetLysSerLysProProProThrTyrGlnSerGln 603
Db 1600 -----CTAAATCTGAAGAT 1614
Qy 604 GluAspLysCysLysProMetSerTyrGlnGlnLysArgGlnLeuSerLeuAspIleAsn 623
Db 1615 GAAGATTAATGCTTAAACCTATGAACTATGATGAGAAAGCAAGTTAGTCTGAATATAAAC 1674
Qy 624 LysLeuProGlyGlnLysLeuGlyArgValAlaHisIleIleGlnSerArgGluProSer 643
Db 1675 AAACCTCCGAGATTAACCTTGGCGAGTAGTTCACATTAATCAATCAAGAGAGCCTTCT 1734
Qy 644 LeuLysAsnSerAsnProAspGluIleGluIleAspPheGlnThrLeuLysProSerThr 663
Db 1735 CTGAGCAATTCCTCAATCTTATGATAGATAGATAGACTTTGAAACACTGAAGACATTAAC 1794
Qy 664 LeuArgGlnLeuGlyAlaLeuCysHisLeuLeuPheAlaGlnGlnLysGlnThrPheLys 683
Db 1795 CTAAAGAAATTTGA-AAAATATGTTTCGGCATG-----TTAAAGAAAGAACCATTAATA 1847

RESULT 5
AAK40172
ID AAK40172 standard; DNA; 3106 BP.
XX
AC AAK40172;
XX
DT 02-JUL-1999 (first entry)
XX
DE WO9904265 Seq ID No: 686.
XX
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
PA (LUDM-) LUDMWG INST CANCER RES.
XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfeundschuh M, Sahin U, Scanlan MJ, Stockert B;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients; used to develop products

```


QY 604 GluAspLysCysLysProMetSerTyrGluGluLysArgGlnLeuSerLeuAspIleAsn 623
 DB 1617 GAAGATTAATCTAAACCTAATGAGATGAGAGAAAAGCAGTTAAGTCTGAATATTAAC 1676
 QY 624 LysLeuProGluLysLysLeuGluArgValValHisIleIleGlnSerArgGluProSer 643
 DB 1677 AAACCTCCCTGGAGATTAACCTTGGCGAGTGTTCACATAATACATCAAGAGCCCTTCT 1736
 QY 644 LeuLysAsnSerAsnProAspGluIleGluIleAspPheGluThrLeuLysProSerThr 663
 DB 1737 CTAGACAAATTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1796
 QY 664 LeuArgGluLeuGluLysLeuGluCysHisLeuLeuPheAlaGluGluLysGluThrPheLys 683
 DB 1797 CTAGAGAAATTAAGA-AAATATATCTTCGGCATG-----TCTAAGAAAGAGACCATTA 1849
 RESULT 6
 AAH18591
 ID AAH18591 standard; cDNA; 1884 BP.
 AC AAH18591;
 XX
 XX 26-JUN-2001 (first entry)
 DE Human cDNA sequence SEQ ID NO:18782.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 PS Claim 8; SEQ ID 18782; 2537bp + CD ROM; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 1884 BP; 492 A; 536 C; 451 G; 405 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1,656-61 Length: 1884
 Score: 1450.00 Matches: 308
 Percent Similarity: 65.78% Conservative: 61
 Best Local Similarity: 54.90% Mismatches: 115
 Query Match: 39.85% Indels: 77
 DB: 22 Gaps: 14
 US-09-700-590A-22 (1-688) x AAH18591 (1-1884)
 QY 16 ValMetGlyAspGlyLeuGluThrSerGlnMetSerThrThrGlnAlaGlnAla----- 33
 DB 330 ATCAGTCGCGCGCATCTGGACACCCCTGCGTGTCTCAGAGAGCACACATCCCTCGGG 389
 QY 34 ---GlnProGlnProAlaAsnAlaAlaSerThrAsnProProProGluThrSerAsn 52
 DB 390 GGCCAATGTGTGCTGTGTGTGCGATTCAAAC-----CCGAGGTGTCTCAAT 437
 QY 53 ProAsnLysProLysArgGlnThrAsnGlnLeuGlnThrLeuLeuArgValValLeuLys 72
 DB 438 CCCAAAAGCCGAGAGGAGTTCACCAACGCTGCAATTAACACACAGGATGATGAAG 497
 QY 73 ThrLeuTrpLysHisGlnPheAlaTrpProPheGlnGlnProValAspAlaValLysLeu 92
 DB 498 GCTGTGTGAAACATCAGTTCGATGCCATTCGCGACCTGTGATCTGTCAAACTG 557
 QY 93 AsnLeuProAspTyrTyrLysIleIleLysThrProMetAspMetGlyThrIleLysLys 112
 DB 558 GGTCTACCGGATTAATCACAAAATTATMAAACGCTTATGACATGGTCTATTAAAGAG 617
 QY 113 ArgLeuGluAsnAsnTyrTyrTrpAsnAlaGlnGluCysAlaGlnAspPheAsnThrMet 132
 DB 618 AGACTTGAAGAACATATTATTATGGCTGCTTCAGAGTGTATGACAAATTATATACATG 677
 QY 133 PheThrAsnCysTyrIleTyrAsnLysProGlyAspAspIleValLeuMetAlaGlnAla 152
 DB 678 TTCACCAACTGTTACATTACAAACAGCCACTGATGATGATGCTTATGACACAAAG 737
 QY 153 LeuGluLysLeuPheLeuGlnLysIleAsnGluLeuProThrGluGluThrIleMet 172
 DB 738 CCGAAGAAAGATATTCTTACAGAAAGTTGCATCAATGCCACAGAAAGAAAGAGCTGTA 797
 QY 173 IleValGlnAlaLysGly-----ArgGly-----ArgGlyArg 183
 DB 798 GTGACCATCTCTAAGAACAGCAAGAAAGGGGGCCCAAGTTGGACGCTCCAGGGCACT 857
 QY 184 LysGluThrGlyThrAlaLysProGlyValSerThrValProAsnThrThrGlnAlaSer 203
 DB 858 GTTACACAGGCCCATCAGAGTGCCTGCTCTCTCTGTTCAACAGACACACCCCTGTAAT 917
 QY 204 ThrProGlnThrGlnThrProGlnProAsnProProProValGlnAlaThrProHis 223
 DB 918 CTTCCACCTGAGATA-----CTACACACGTCTCTCAACATTCCTCCAC 959
 QY 224 ProPheProAlaValThrProAspLeuIleValGlnThrProValMetThrValValPro 243
 DB 960 CCA-----TCAGCATTCCTCTCCACTTC----- 986
 QY 244 ProGlnProLeuGlnThr-----ProProProValProProGlnProGlnProPro-Pr 261
 DB 987 ---AGTCTTGTCACTCTGTGAGACCCCGCTCTTGTGTTACGTGACGCTCTCCAGCC 1043
 QY 261 oAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThrProGlnPr 281
 DB 1043


```

QY 128 AspPheAsnThrMetPheThrAsnCysTyrIleTyrAsnLysProGlyAspAspIleVal 147
DB 1272 GACCTCAATCGATGTTCAACAAATTGCTATGCTACAAACAGCCGCGAGATGATAGT 1331
QY 148 LeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLysIleAsnGluLeuProThrGlu 167
DB 1332 GTTATGCCCCAGCGCTCGAAGAGTCTTCTCCAGAAATCGAATGATCGCTTAAGAG 1391
QY 168 GluThrGluIleMetIleValGlnAlaLysGlyArgGlyArgGlyArgGlyArgGly 187
DB 1392 GACCTGMACTGAGCGCGGTTCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1451
QY 188 ThrAlaLys-----ProGlyValSerThrValProAsnThrThrGlnAla 202
DB 1452 ACCCCCAAGCATGTCGCGGTGCGCTGAGACATCCACCGTTCTGCTATCATCTCGCGCA 1511
QY 203 SerThrProProGlnThrGlnThrProGlnProAsn-ProProProValGlnAlaThrPr 222
DB 1512 GCGATT-----ACGAGCGGACGAGTAGTGGCTCCACAAAGTGTGATGCTGCC 1562
QY 222 ohisProPhe-----ProAlaVal-----ThrProAspLeuIleVal 234
DB 1563 TCATCCGACACAAAGTCCGCTCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1622
QY 234 IGIThrProValMetThrValValProProGlnProLeuGlnThrProProValPr 254
DB 1623 ACACCGGAACTCAGCGCGGTTCCGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1682
QY 254 oPro-----GlnProGlnProProProAlaProAlaProGlnProValGln 269
DB 1683 GGCATGGGCGGACGCGTTTCATCGACGCGCGCGGT-GCACCGTCCATACACCGATTAG 1741
QY 269 nserHisProPro----- 273
DB 1742 CACAAATGCCACCGCACACGATACCCGAGCACCAATAGACAGACAGACAGATGCTGG 1801
QY 273 ----- 273
DB 1802 CGCGTTGGTGGCGCAAGTGCAGCTGGAGCCAAATCCCAATGTGCGGCTCGATGGCTAG 1861
QY 273 ----- 273
DB 1862 TCTTTCGAATGCGGCGCAACGCGCGCTTATCCCGGTGCGCCGCGCGACAGCGCGTCAA 1921
QY 273 ----- 273
DB 1922 TAGCTCTGCTTCTAGATGGCAGTACAGCCGAGTAGAGCGGCGAGCAGACGCGCTGC 1981
QY 273 ----- 273
DB 1982 GCGGCGGCGCGCGCGGAGAGGTGACGCGGAGCGCGTGAAGAGCAGAGCAATATAC 2041
QY 273 ----- 273
DB 2042 AGCAGTTGCGGTCAATGCGCGCTAACGCGTTCAAGCGCTTATGATATCGGCGGTGAGCGT 2101
QY 274 -----IleIleAlaAlaThrPro-----GlnProValLysThrLysLysGlyValLysAr 290
DB 2102 CGGAGTGGAGCGCGGTATACCGCTCAGCAGCGCGCAAAATTAAGAGGAGGTGCAACG 2161
QY 290 GlyAlaAspThrThrThrProThrThrIleAspProIleIleGluProProSerLeuPr 310
DB 2162 GAGGCGGAGACAAACGCGCGAGCGCC-----AATGCTTTGAATCCCGTACACGCA 2215
QY 310 oProGluProLysThrThrLysLeuGlyGlnArgGlyLysSerArgProValLysPr 330
DB 2216 AATGAGCTCCAAATCGCGCAAGATTGGACGCGCGGAGATCGAATCGT----- 2264
QY 330 oProLysLysAspVal-----ProAspSerGlnGlnHisProAlaPro----- 344
DB 2265 -----CAGGATCTTACATTCCAGGCGTGGAGATACATATGCGCGGTAGGCGTCTC 2317

```

```

QY 345 -----G1 345
DB 2318 CGGAGTCCCGGACTTGGCGGTCTACTGCGCGCGCGCTGGTGTGTTGCGTGCCCA 2377
QY 345 uLysSerSerLysValSerGlnGlnLeuLysCysCysSerGlyIleLeuLysGluMetPr 365
DB 2378 GAACAAAGAGAACGTCGATGCGCTCAAGTGTCTCAACGAATCTTCAAGAGCTCTT 2437
QY 365 eAlaLysLysHisAlaAlaTyrAlaThrProPheTyrLysProValAspValGluAlaLe 385
DB 2438 CTCGAAGAAGCACTCGGCGCTATGCTTGGCCATTCTACAGCCAGTGTGACGCGGAATGCT 2497
QY 385 uGlyLeuHisAspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSe 405
DB 2498 TGGCTTCGACTGATACACAGCATCATCAAGAAACCAATGATCTGGCAGCTCAAGCG 2557
QY 405 LysLeuGlnAlaArgGluTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMe 425
DB 2558 GAAATGAGCAATCGGAGTACAGAGCGCGCGGAATTTCCCGCGAGTGGATTTAAT 2617
QY 425 cPheSerAsnCysTyrLysTyrAsnProProAspHisGluValAlaAlaMetAlaArgly 445
DB 2618 ATTCACCAACTGCTACAGTACATCGGCAAGATCATGATGTTGTGGCATGGGTGCGCA 2677
QY 445 sLeuGlnAspValPheGluMetArgPheAlaLysMetProAspGluProGluLysProVa 465
DB 2678 GCTGCAAGAGCTATTGAGATGGCTATGCCAAATCCCGATGAGCCGGTAGCCCAATGC 2737
QY 465 lVal-lAlaValSerSerProAla----- 472
DB 2738 GGGCCACATACAGGACATGCGCGGAGATGTCATGCGCATGGCCATGGCCATCGGCGCA 2797
QY 473 -----ValProProThrLysValAlaLysProProSerSerSerSers 489
DB 2798 TGCTACAGGACATGTCATGATGTCACGAGATACGCGGCTCTCTCTCACTCAAGAC-GATG 2856
QY 489 eSerAspSerSerSerAspSerSerSerSerThrAsp-----AspSerGluGlu 506
DB 2857 CCAGCATTCGTCACGCGAGGACTCCAGCGATACCAAAACGAATCCCACTCGATGAGG 2916
QY 506 lAlaArgAlaGlnArgLeuAlaGluLeuGlnGlnLeuLysAlaValHisGluGlnLeuA 526
DB 2917 AGCGAGCGCTAGGCTGAAGATGCTGAGTCCAACTGCTCGCTGTCAGAGAGAAATCC 2976
QY 526 lAlaLeuSerGlnProGlnGlnAsnLysProLysLysLysGlyLysAspLysLysGluL 546
DB 2977 GTAAGCTGTCAAGAGAGGCTCCGCCAAAGAGGCGAAGAAAACTCAGAGAAAAAGA 3036
QY 546 ySLysLys----- 548
DB 3037 AAAAGTCATTAAGCGCGGATCAGGCTGTGTTGCGCTTCGACCATGTGACGCCACGG 3096
QY 548 ----- 548
DB 3097 GCGGCGGTGCMAATGCTGCGGAGAGAGCGGTCGCCGATCAGCGCGCATGGAGCGTTT 3156
QY 548 ----- 548
DB 3157 CGGTCCAGGCGGTGTGGGTCTTGGGTCCGGTGGAGCGGCGCGCTAATCTCAAG 3216
QY 548 ----- 548
DB 3217 CGGTGTCGGTGCTCATTTGTTGGCCATGGCGAGCGGCGCTCGCAGAGAGGCGTTCCA 3276
QY 549 -----GluLysHisLysArgLysGluGluValGlu----- 558
DB 3277 AATGAGTTCGTTGACACACCGATTCAAGACGTGCGCATGCGCTTTAGCCAGATGGCGG 3336
QY 558 ----- 558
DB 3337 GCGGCGGTCCCGCGCGGTGCTGGCTTGTGTCGCGGTGTGACAGCAGAGCATCGT 3396
QY 558 ----- 558

```


QY 65 TyrLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAlaTrpProPheGln 84
 Db 229 TACCTCGCGCAGGTAGTGTGAAAGGCTGTGTGAAACATCATGCTCCCATGGCCATTCGG 288
 QY 85 GlnProValAspAlaValLysLeuAsnLeuProAsp----- 96
 Db 289 CAGCCTGTGATGTCTGTCAAACTGGGCTTACCGAT-TCCCACTCGGGTTGGAGAGAGAC 347
 QY 96 ----- 96
 Db 348 CACGGTGGCAAAATCTTAGCTTCTTCCCTTCATCGACCCATGATGATGCCAGCC 407
 QY 97 -----TyrTrpLysIleIleLysThrProMetAspMetGlyThrIleLysArg 113
 Db 408 CCAGAGGATTAATCATCAAAATTAATAAAGAGCTATGAGATGGGTACTATTAGAGAGAGA 467
 QY 114 LeuGluAsnAsnTrpTyrTrpAsnAlaGlnGluCysIleGlnAspPheAsnThrMetPhe 133
 Db 468 CTTGAAAACAAATTATTATGGGCTGCTCAGAGTGTATGCMAATTTTAAATACCATGTC 527
 QY 134 ThrAsnCysTyrIleTyrAsnLysProGlyAspAspIleValLeuMetAlaGluAlaLeu 153
 Db 528 ACCAATGTTACTTTCATTCACAGAGCCCATGATGATATTGTCTTAATGGCACAAGCCTG 587
 QY 154 GluLysLeuPheLeuGlnLysIleAsnGluLeuProThrGluThrGluIleMetIle 173
 Db 588 GAAAGATTAATCTTACAGAAAGTTGCATCATGCTCAAGAGAAACAAGAGCTGTAGTG 647
 QY 174 ValGlnAlaLysGly-----ArgGly-----ArgGlyArgLys 184
 Db 648 ACCATCCCTAAGAAACAGCCACAAGAGGGGGCCAGTTGGCAGCGCTCCAGGGCAGTGT 707
 QY 185 GluThrGlyThrAlaLysProGlyValSerThrValProAsnThrThrGlnAlaSerThr 204
 Db 708 ACCAGTGGCCATAGGTGCTGCTGCTCTCTGTGTGTACACACAGCCCTGTATCTCT 767
 QY 205 ProProGlnThrGlnThrProGlnProAsnProProProValGlnAlaThrProHisPro 224
 Db 768 CCACCTGAGATA-----CTTACCACTGTCTCTCAACATTTCCCAACCA 809
 QY 225 PheProAlaValThrProAspLeuIleValGlnThrProValMetThrValValProPro 244
 Db 810 -----TCAGTCATTTC-TCTCCACTTCTCAAGCTTGCAC- 844
 QY 245 GlnProLeuGlnThrProProValProProGlnProGlnProProAlaProAla 264
 Db 845 -----TCTGCTGAGCCCGCTCTCTGTGTGTACTGACAGCTCTCCAGCC- 889
 QY 265 ProGlnProValGlnSerHisProProIleIleAlaAlaThrProGlnProValLysThr 284
 Db 890 -----CAGCCCTTGGCAAG 904
 QY 285 LysLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle----- 300
 Db 905 AAAAAAGGCGT-AAGCGGAAGCAGATACTACCAACCCCTACACCTACAGCCATCTGGCT 963
 QY 301 -----AspProIleHisGlnProProSerLeuProProGlnProLysThrThrLysLeu 318
 Db 964 CTTGTTCTTCCACTAGCTCTCTGGAGTCTT-----GAGCCTAAGGAGCAGACCCCTT 1017
 QY 319 ---GlyGlnArgArgLysSerSerArgProValLysProProLysLysAspValProAsp 337
 Db 1018 CCCCCATGTGTAGAGAGATGTGCTGCCCATCAGGCCCCCAAGCAAAGCTTGGCTGAC 1077
 QY 338 SerGlnGlnHisProAlaProGluLysSerSerLysValSerGluGlnLeuLysCysCys 357
 Db 1078 TCTCACACACACACACAGAGCTTAAGAAAG-AAAAGCTTCAAGACGTTAAACAATTGC 1136
 QY 358 SerGlyIleLeuLysGluMetPheAlaLysValHisAlaAlaTyrAlaTrpProPheTyr 377
 Db 1137 AATGGAATT-TTGAAGGAGTTACTCTCTAAGAGCATGCTCCATAGTGTGGCTTTCTAT 1195
 QY 378 LysProValAspValGluAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisPro 397

Db 1196 AAACCGTGAATGCTTCTCTCACTTGCGCTGCATGACTACCATGACATCATTAAGCCCC 1255
 QY 398 MetAspMetSerThrIleLysSerLysLeuGluAlaArgGluTyrArgAspAlaGlnGlu 417
 Db 1256 ATGACCTCAGACACTCTCTCAAGCGAAGATGAGAGAACCGATTATCCGGATGACAGAG 1315
 QY 418 PheGlyAlaAspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHis 437
 Db 1316 TTGTGCTGTAGTACGGCTTATGTC-TCCAACTGCTATAGTACAT-GCCCCAGATCAC 1373
 QY 438 GluValValAlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMet 457
 Db 1374 GATGTGTGGCAATGGCACAGAACTTACAGATGAT-TTACTTCGTTATGCGCAAGATG 1430
 QY 458 ProAspGluProGluGluProValValAlaValSerSerProAlaValProProPro 476
 Db 1431 CCAGATGAACCACTAGAACCAAGGCTTTTACAGATCTTACTGCTCCATGCCCCCTGCTG 1490
 QY 477 ThrLysValValAlaProProSerSerSerAspSerSerSerSerSerSerSer 494
 Db 1491 GCCAATCGTCTGTACAGATGCTCCAGTGAAGAAAGTACGAGAGCTCTCTGAGGA 1550
 QY 495 -----AspSerAspSerSerThrAspAspSerGlu 504
 Db 1551 GAGAGAGAGAAATGAGAGAGACGAGAGAAAGAAAGAGTGAAGCTCAGACTCAGAG 1610
 QY 505 GluGluArgAlaGlnArgLeuAlaGluLeuGlnGluLeuLysAlaValHisGluGln 524
 Db 1611 GAAAGAAAGGCTCATGCTTATGACAGAACTTACAGAAACAGCTTCCGCAATGACAA 1670
 QY 525 LeuAlaAlaLeuSerGlnProGlnGlnAsnLysProLysLysGluLysAspLysLys 544
 Db 1671 CTGGCTGCTGTGTCCAGAGGTCCGATATCCAAAGCCCAAGAGGAAAAAGAAAAA 1730
 RESULT 9
 ABV25740
 ID ABV25740 standard; cDNA, 1862 BP.
 XX
 AC ABV25740;
 XX
 DT 16-SBP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 25731.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 XX
 PR 16-MAR-2000; 2000US-189862P.
 XX
 PR 25-MAY-2000; 2000US-207454P.
 XX
 PR 09-JUN-2000; 2000US-211314P.
 XX
 PR 18-JUL-2000; 2000US-219007P.
 XX
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer, useful
 XX for detecting presence of prostate cancer, stage of prostate cancer -

PS Claim 1, Page 5167-5168, 11750bp; English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 1862 BP, 502 A, 514 C, 483 G, 363 T, 0 other;

Alignment Scores:

Pred. No.:	1,22e-47	Length:	1862
Score:	1161.50	Matches:	274
Percent Similarity:	59.40%	Conservative:	64
Best Local Similarity:	48.15%	Mismatches:	131
Query Match:	31.92%	Indels:	100
DB:	23	Gaps:	17

US-09-700-590A-22 (1-688) X ABV25740 (1-1862)

QY 126 IIEGLINASPHEANTHMETPHERTHASNCYSTYRILETYRAENLYSPROGLYASPH 145
DB 16 ATTAATACTTGACATACATGTTCCACCAACTGTGCTTACGACAAATATCATCAGTGGC 75
QY 146 IIEVALLEUMETALAGUALALEUGLULYSLEUHELEUGLULYSLEIASEGLULEUPRO 165
DB 76 TTGTTGTAATGTGTAACAAACGCTGGAAGATATTCCTACAGAAAGTTGCAATCAATGCCA 135
QY 166 Thcglugluthrgluilemetilevalglinalals 177
DB 136 CAAAGAAACAAAGAGCTGTGTAACCATCCCTAAGAACACACAGAGGGGCCAATT 195
QY 178 GLIARGELY-ARGGLYARGLYSGIUTHRGLYTHRALALSPROGLYVALSERTHVALPR 197
DB 196 GCGAGCGCTCGGGGAGTGTACACAGTCCATCAGGTGCTCGCTCTCTGTGTCTC 255
QY 197 OASNTHTHGLINLASERTHPRPROGLINTHGLINTHPRPROGLINPROANPROPROR 217
DB 256 ACACACAGCCCTCTATCTCTCTCCACCTGAGATA- 297
QY 217 OVALGLINALATHRPHIOPHOPHEPROALVALTHRPROASPLEULEVALGLINTHRPR 237
DB 298 TGTCTCAACATTCCTCCCA- 336
QY 237 OVALMETTHRVALVALPROGLINPROLEUGLINTHRPROPROVALPROPROGLINPR 257
DB 337 ACTTCTCAAGCTCTTGAC- 384
QY 257 OGLINPROPROALAPROALAPROGLINPROVALGLINSERHISPROFOLLEILEALAI 277
DB 385 TGCAGCTCTCTCAGC- 400
QY 277 aTHPRGLINPROVALYSLYSLYGLYVALYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 297
DB 401 -CAGCCCTTGCCAAAGAAAGGCTTAAGCGGAAGAGATACACACCC 453
QY 297 OTHTHRILE- 311
DB 454 TACACCTACAGCACTTGGCTCTGTTCTCAGCTAGCCCTCTGGAGAGCTT- 508
QY 311 OGLUPROLYSTHRTHTYSLYSLY-GLYGLINARGLYSLYSLYSLYSLYSLYSLYSLYSLY 330
DB 509 -GAGCTTAAGGACAGCAGCTTCCCTATGCTAGAGAGAGTGTGCTGCTCCATCAAGCC 567

QY 330 OPROLYSLYASPVALPROASPSERTGLINHSIPIROALAPROGLULYSERSLYSLY 350
DB 568 CCCACCAAGACTCTCTGACTCTGACAAACACACACAGCTTAAAGAAAGAGAGCT 627
QY 350 ISEGLUGLINTLEUYCYCSERGLYILELEUYSGIUMETPHEALALYSLYSHISA 370
DB 628 TTGAGAACACTTAAACATTCGACATGACATTTGAAGAGTTACTCTCAAGAGATCG 687
QY 370 AALATYRALATTPROPHETIYSLYSPROVALASPVALGLUALALEUGLYLEUHSIPRY 390
DB 688 TGCCTATGCTTGCTCTTCTATTAACAGAGATGCTTGCTGACCTTGCTGATGACTA 747
QY 390 IYSAEPILILELYHISIPROMETASPMETSERTHRIELYS- 404
DB 748 CCATGACATCATTAACACCCACATGACCTGACACTGTCACTTGTGATGCTGCTCC 807
QY 405 -SERLYSLEUGLUALIARGGLUTYRARGSPALAGIINGLUPHEGLYALASPV 422
DB 808 TTCTGACGCGAAGATGGAACACCTGATTAACCGGATGACAGAGCTTGTCTGATG 867
QY 422 ALARGLEUMETPHEASERASNCYSTYRILYSTRANBPROPROASPHISGLUALVALALAM 442
DB 868 TACGGCTTATGTTCTCAACTGCTATAGTACATCCCAATCAGATGTTGTGGCAA 927
QY 442 ECLALARGLYSLEUGINASPVALPHEGLUMETARGPHEALALYSLMETPROASGLUPROG 462
DB 928 TGGCAGCAAAAGCTTACAGAGATTTGAGTTCCGTATATGCCAAGATGCCAGATGAACAC 987
QY 462 IUGLUPROVALVALALVALSERSEPROVALVALPROPRO-THLYVALVALA 481
DB 988 TAGAACACAGGCTTTTACAGCTCTTACTCTACATGCCCTTGCTTGCCAAATCGCTT 1047
QY 481 LAPROPSERSESERSESERSESERSESERSESERSE- 494
DB 1048 CAGAGCTCTCAGTGAAGAAAGTAGCAGTAGAGCTCTCTGAGAAAGAGAGAGAG 1107
QY 495 - 509
DB 1108 ATGAGAGAGACGAGAGAAAGAGAGTAGTAAAGCTTCAAGCTCAGAGAGAAAGGCTC 1167
QY 509 IARGLEUALIAGLULEUGLUGLINTLEUYSLAVALHISGLUGLINTLEUALALALEUS 529
DB 1168 ATGCTTAAAGCAACTTACAGAAACGCTTGGCAGATACATGAACACTGCTCTCTGT 1227
QY 529 ERGLINPROGLINASLYSPROLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 548
DB 1228 CCCAGGCTCAATATCAAGCCCAAGAGAAA-AGAGAGAAAAAGAGAAAAAGAGA 1284
QY 549 -GLULYSHISLYSEHLYSGIUGLULVALIUGLULASRLYSLYSLYSLYSLYSLYSLYSLY 564
DB 1285 AACGAGAGACAGAGAAACATGAGGCGAGCTGGGCGAGTGAAGATGAGACAAAGGCTTA 1344
QY 564 YSALALYSGIUPRO-PROLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 583
DB 1345 GGGACACCCCGCCCACTTCAACTTAAGAGTCAAGAAAGCAAGTGGCAGTGGGGTGGCA 1404
QY 583 ALSERTLYSGIUPRO- 591
DB 1405 GTGCTGCTTAAGGCTCTTGCTTGACCTTGTGAGACCTTGTGAGGAAGTGGCACCAGCTCCC 1464
QY 591 ELYSERTLYS-PRO-PROTHRTYRGLUSERGLUGLUGLU 604
DB 1465 AAAAAGCCCAAGACAGCCCACTGCTGCTTCAAGATTATGATTAGAGAGAGAG 1524
QY 605 ASPLYSCYSLYSPROMETSETERTYRGLUGLULYSLYSLYSLYSLYSLYSLYSLYSLY 624
DB 1525 GAAAGAGAGCCCAATGATGATGAGAGAGGAGGAGCTGAGCTGAGCATCAACAAA 1584
QY 625 LEUPROGLYGLULYSLYSLY 631
DB 1585 TTACTGAGGCCCCAAGAGGA 1605

RESULT 10
 AAS03029
 ID AAS03029 standard; cDNA; 667 BP.
 XX
 XX AAS03029;
 AC
 XX
 DT 29-AUG-2001 (first entry)
 XX
 XX Human diagnostic and therapeutic (dthp) cDNA sequence #18.
 DE
 XX Human diagnostic and therapeutic molecule; dthp; gene therapy;
 KW thalassemia; cardiovascular disorder; cell proliferative disorder;
 KW cancer; neurodegenerative disorder; autoimmune disorder;
 KW infectious disorder; inflammatory disorder; developmental disorder;
 KW Incyte ID number 2432679dec; intracellular signalling molecule; ss.
 OS Homo sapiens.
 XX
 XX WO200121836-A2.
 PN
 XX
 PD 29-MAR-2001.
 XX
 PF 19-SEP-2000; 2000MO-US25643.
 XX
 XX 23-SEP-1999; 99US-0155760.
 PR 24-SEP-1999; 99US-0155939.
 PR 24-SEP-1999; 99US-0156294.
 PR 28-SEP-1999; 99US-0156565.
 PR 28-SEP-1999; 99US-0156624.
 PR 28-SEP-1999; 99US-0156625.
 PR 24-NOV-1999; 99US-0167410.
 PR 24-NOV-1999; 99US-0167453.
 PR 24-NOV-1999; 99US-0167517.
 PR 24-NOV-1999; 99US-0167520.
 PR 24-NOV-1999; 99US-0167542.
 PR 29-NOV-1999; 99US-0167943.
 PR 29-NOV-1999; 99US-0167945.
 PR 30-NOV-1999; 99US-0168197.
 PR 30-NOV-1999; 99US-0168265.
 PR 30-NOV-1999; 99US-0168429.
 PR 30-NOV-1999; 99US-0168432.
 PR 01-DEC-1999; 99US-0168468.
 PR 01-DEC-1999; 99US-0168599.
 XX
 XX (INCYTE-) INCYTE GENOMICS INC.
 PA
 XX Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;
 PI Bratcher SR, Dufour GE, Cohen HU, Rosen BH, Shah P, Chalup MS;
 PI Hillman JL, Jones AL, Yu JY, Greenwalt LB, Panzer SR;
 PI Roseberry AM, Wright RJ, Chen W, Liu TF, Yap PE, Stockdreher TK;
 PI Ameshey S, Fong WT;
 PI
 XX MPI; 2001-281607/29.
 XX
 XX Novel diagnostic and therapeutic polynucleotides, used in disease
 PT diagnosis and for gene therapy of conditions such as cancer and
 PT thalassemia
 PT
 XX
 XX Claim 1; Page 263; 299pp; English.
 PS
 XX
 XX The present sequence for human diagnostic and therapeutic (dthp) cDNA
 CC sequence #18 is 1 of 71 (AAS03012-AAS03082) novel sequences described
 CC in the invention. The present sequence (Incyte ID No: 2432679dec)
 CC encodes an intracellular signalling molecule. The dthp polynucleotides
 CC may be used to diagnose a condition disease or disorder associated with
 CC human molecules. They can be used to identify the presence of similar
 CC nucleic acids. Dthp polynucleotides may be used to generate hybridisation
 CC probes for use in chromosomal mapping. Polypeptides (DTHP) encoded by
 CC dthp are used to screen for molecules which bind to them and modulate
 CC their activity. Dthp polynucleotides can be used for gene therapy of
 CC disorders such as severe combined immunodeficiency syndrome (SCID),
 CC cystic fibrosis, thalassemia, haemophilia resulting from Factor VIII
 CC or IX deficiencies, cardiovascular disorders e.g familial

CC hypercholesterolaemia (RH), cell proliferative disorders e.g. cancers,
 CC neurodegenerative disorders, autoimmune/inflammatory disorders,
 CC infectious disorders and developmental disorders. The antibodies can be
 CC used to analyse protein expression levels.
 CC
 XX
 XX SQ Sequence 667 BP; 229 A; 157 C; 149 G; 132 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,23e-39 Length: 667
 Score: 981.00 Matches: 186
 Percent Similarity: 98.95% Conservative: 3
 Best Local Similarity: 97.38% Mismatches: 1
 Query Match: 26.96% Indels: 1
 DB: 22 Gaps: 0
 US-09-700-590A-22 (1-668) x AAS03029 (1-667)
 QY 1 MetSerAlaGluSerGlyPro-GlyThrArgLeuArgAsnLeuProValMetGlyAspG 20
 DB 22 ATGTCTGCGGAGAGCGCCCTGGGGACGAGATTGAGAAATCTGCCAGTATGGGGATGG 81
 QY 20 YLeuGluThrSerGluMetSerThrThrGlnArgGlnArgGlnProGlnProAlaAsnAl 40
 DB 82 ACTGAAACTTCCCAATGCTTAACAACAGGCCAGGCCAACCCAGCCAGCCAAACGC 141
 QY 40 aAlaSerThrAsnProProProGluThrSerAsnProAsnLysProLysArgGlnH 60
 DB 142 AGCCACACCAACCAACCCCGCCGACAGACCTCAACCTTACAAACCAAGGAGCAGC 201
 QY 60 rAaGlnLeuGlnThrLeuLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAl 80
 DB 202 CAACCAACTGCATATACCTGCTCAGAGTGTGCTCAAGACACTATGAAACACCAAGTTTCG 261
 QY 80 aATPProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTrpTrpLysH 100
 DB 262 ATGGCTTTTCAGACAGCTGTGATCCGTCAGCTGAACCTCCGATTACTATAGAT 321
 QY 100 eLeuLysThrProMetAspMetGlyThrLysLysAspGluGluAsnAntyTrpTr 120
 DB 322 CATTAAACCCCTATGATGATGGAACATTAAGAAGCCCTTGAAACCAACATTTACG 381
 QY 120 pAaAlaGlnGluCysLeuGlnAspPheAsnThrMetPheThrAsnCysTrpLysTrAs 140
 DB 382 GAATGCTCAGGAATGTATCAGACCTTCAACTATGTTTACAAATGTTTACATCAACAA 441
 QY 140 nLysProGlyAspAspAlaValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGln 160
 DB 442 CAAGCTTGAGATGACATGATCTTAATGACAGAGCTTGGAAGAAAGCTCTTCTGCAAAA 501
 QY 160 sLeaGlnGluLeuProThrGluGluThrGluLysMetIleValAlaAlaLysGlyArgG 180
 DB 502 AATTAATGAGCTTACCCACAGAGAAACCAAGATCATGATGTCAGGCAAAAGAGAGG 561
 QY 180 yArgGlyArgGlyGluThrGlyThrAlaLys 190
 DB 562 ACCTGGAGAAAAGAAAGTGGGTTATCAAGG 592
 RESULT 11
 AAS26193
 ID AAS26193 standard; cDNA; 746 BP.
 XX
 XX AAS26193;
 AC
 XX
 DT 07-NOV-2001 (first entry)
 XX
 XX Human cDNA encoding a novel secreted protein, Seq ID 372.
 DE
 XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;
 KW cytototoxic; cardiac; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; varicide; fungicide; ophthalmological;
 KW vulnary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischemia; angiogenesis;

KM nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KM corneal infection; wound healing; epithelial cell proliferation;
KM skin ageing; food additive; preservative; antiproliferative.

OS Homo sapiens.

PN WO200155322-A2.

PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01341.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225256.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 23-AUG-2000; 2000US-0227182.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239335.

PR 13-OCT-2000; 2000US-0239337.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249219.

PR 17-NOV-2000; 2000US-0249224.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249246.

PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.

PR 06-DEC-2000; 2000US-0256719.

PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX MPI; 2001-488783/53.
 DR P-PSDB; AAU16206.
 XX
 PT New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 PS
 PS Claim 1; SEQ ID No 372; 980bp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence encodes a novel secreted protein of the invention.

Alignment Scores:
 Score: 3.05e-25 Length: 746
 Percent Similarity: 688.00 Matches: 144
 Best Local Similarity: 66.40% Conservative: 22
 Query Match: 57.60% Mismatches: 61
 DB: 18.91% Indels: 24
 Gaps: 6

US-09-700-590A-22 (1-688) x AAS26193 (1-746)

QY 51 SerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGlnTyrLeuLeuArgValVal 70
 DB 3 TCCAAACCCGAGCAAGCCGCGGCAAGCAACCAAGCTGACGTACAGAAATGTGGTG 62
 QY 71 LeuLysThrLeuTyrLysHisGlnPheAlaTyrProPheGlnGlnProValAspAlaVal 90
 DB 63 GTGAAGACGCTCTGGAACACCACTTCGCTGCGCTTCTTCAAGCCCGTGAAGCAATC 122
 QY 91 LysLeuAsnLeuProAspTyrTyrLysIleIleLysThrProMetAspMetGlyThrIle 110
 DB 123 AAATTTGAACCTGGCGGATTTATCATTAATAATTAATAAACCAATGATATGGGACATT 182
 QY 111 LysLysArgLeuGlnAsnAsnTyrTyrTyrAsnAlaGlnGluCysIleGlnAspPheAsn 130
 DB 183 AAGAGAGACTAGAAAAATATTATTATTTGAGTCAGACCGAATATATGAGACTTCAAC 242
 QY 131 ThrMetPheThrAsnGlyTyrIleTyrAsnLysProGlyAspAspIleValLeuMetAla 150
 DB 243 ACCATGTTTACAAATGTTTACATTATTAACAGCCACAGATGACATAGGCTATATGACC 302
 QY 151 GlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGlnLeuProThrGlnGluThrGln 170

DB 303 CAAGCTTTAGAGAAATTTTCTTACAAAAAGTGCCAGATGCCCAAGAGCAAGTTGAA 362
 QY 171 IleMetIleValGlnAlaLysGlyArgGlyArgGlyArgGlyArgGlyThrAlaLys 190
 DB 363 ATTATTATACCCCTGCTCCAAAGGAGCAAGAGTGGAGCGCT-GGGGAGCCCAAGAC 421
 QY 191 ProGlyValSerThrValProAsnThrThrGlnAlaSerThrProGlnThrGlnThr 210
 DB 422 GCAGGTACACAGCAAGTGGGCGCGCTCTCTGTCTCC-----CCAGCGACC 469
 QY 211 Pro---GlnProAsnProProProValGlnAlaThrProHisProPheProAlaValThr 229
 DB 470 CCTTTTCAAGCGCTGCCCGCCACCGTCTCCACAGCGCC----- 508
 QY 230 ProAspLeuIleValGlnThrProValMetThrValValProProGlnProLeuGlnThr 249
 DB 509 -----GTATTCGCTGCACCCCTGTATCAACATACATCTCAAAC-----GTCACG 553
 QY 250 ProProProValProProGlnProGlnProProProProAlaProAlaProGlnProValGln 269
 DB 554 TCGGTCCAGTCCCGCCAGCTGCCCGCCCTCTCT-----CCTGCCACA 598
 QY 270 SerHisProProIleIleAlaAlaThrProGlnProValLysThrLysGlyValLys 289
 DB 599 CCATCGCTCCCGGTGGTCCCTCTCTACGCCA---CTGTCTCAAGAAANAGCGTTGAA 655
 QY 290 ArgLysAlaAspThrThrThrProThrThr 299
 DB 656 CGGAAGCAACACACACACTCCACGACT 685

RESULT 12
 AAS26606
 ID AAS26606 standard; cDNA; 424 BP.
 XX
 AC AAS26606;
 XX
 DT 07-NOV-2001 (first entry)
 DE Human cDNA encoding a novel secreted protein, Seq ID 785.
 XX
 KW Human; immunosuppressive; antiarthritic; ss; antineumatic;
 KW cytoskeletal; cardiac; vasotrophic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin aging; food additive; preservative; antiproliferative.
 OS Homo sapiens.
 XX
 PF 17-JAN-2001; 2001W0-US01341.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-020515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR	4-JUL-2000	2000US-0218829
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226281
PR	22-AUG-2000	2000US-0226286
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227189
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231415
PR	08-SEP-2000	2000US-0233081
PR	08-SEP-2000	2000US-0233082
PR	12-SEP-2000	2000US-0233968
PR	14-SEP-2000	2000US-0233997
PR	14-SEP-2000	2000US-0233998
PR	14-SEP-2000	2000US-0234399
PR	14-SEP-2000	2000US-0234400
PR	14-SEP-2000	2000US-0234531
PR	14-SEP-2000	2000US-0234532
PR	14-SEP-2000	2000US-0234533
PR	14-SEP-2000	2000US-0234534
PR	14-SEP-2000	2000US-0234535
PR	14-SEP-2000	2000US-0234536
PR	14-SEP-2000	2000US-0234537
PR	25-SEP-2000	2000US-0234597
PR	25-SEP-2000	2000US-0234598
PR	25-SEP-2000	2000US-0234599
PR	26-SEP-2000	2000US-0234584
PR	26-SEP-2000	2000US-0234585
PR	26-SEP-2000	2000US-0234586
PR	26-SEP-2000	2000US-0234587
PR	26-SEP-2000	2000US-0234588
PR	26-SEP-2000	2000US-0234589
PR	26-SEP-2000	2000US-0234590
PR	26-SEP-2000	2000US-0234591
PR	26-SEP-2000	2000US-0234592
PR	26-SEP-2000	2000US-0234593
PR	26-SEP-2000	2000US-0234594
PR	26-SEP-2000	2000US-0234595
PR	26-SEP-2000	2000US-0234596
PR	26-SEP-2000	2000US-0234597
PR	26-SEP-2000	2000US-0234598
PR	26-SEP-2000	2000US-0234599
PR	26-SEP-2000	2000US-0234600
PR	26-SEP-2000	2000US-0234601
PR	26-SEP-2000	2000US-0234602
PR	26-SEP-2000	2000US-0234603
PR	26-SEP-2000	2000US-0234604
PR	26-SEP-2000	2000US-0234605
PR	26-SEP-2000	2000US-0234606
PR	26-SEP-2000	2000US-0234607
PR	26-SEP-2000	2000US-0234608
PR	26-SEP-2000	2000US-0234609
PR	26-SEP-2000	2000US-0234610
PR	26-SEP-2000	2000US-0234611
PR	26-SEP-2000	2000US-0234612
PR	26-SEP-2000	2000US-0234613
PR	26-SEP-2000	2000US-0234614
PR	26-SEP-2000	2000US-0234615
PR	26-SEP-2000	2000US-0234616
PR	26-SEP-2000	2000US-0234617
PR	26-SEP-2000	2000US-0234618
PR	26-SEP-2000	2000US-0234619
PR	26-SEP-2000	2000US-0234620
PR	26-SEP-2000	2000US-0234621
PR	26-SEP-2000	2000US-0234622
PR	26-SEP-2000	2000US-0234623
PR	26-SEP-2000	2000US-0234624
PR	26-SEP-2000	2000US-0234625
PR	26-SEP-2000	2000US-0234626
PR	26-SEP-2000	2000US-0234627
PR	26-SEP-2000	2000US-0234628
PR	26-SEP-2000	2000US-0234629
PR	26-SEP-2000	2000US-0234630
PR	26-SEP-2000	2000US-0234631
PR		

PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.*
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	17-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249219.
PR	17-NOV-2000;	2000US-0249220.
PR	17-NOV-2000;	2000US-0249221.
PR	17-NOV-2000;	2000US-0249222.
PR	17-NOV-2000;	2000US-0249223.
PR	17-NOV-2000;	2000US-0249224.
PR	17-NOV-2000;	2000US-0249225.
PR	17-NOV-2000;	2000US-0249226.
PR	17-NOV-2000;	2000US-0249227.
PR	17-NOV-2000;	2000US-0249228.
PR	17-NOV-2000;	2000US-0249229.
PR	17-NOV-2000;	2000US-0249230.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251865.
PR	08-DEC-2000;	2000US-0251866.
PR	08-DEC-2000;	2000US-0251867.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
P1	WPI; 2001-488783/53.	
XX	P-PDB; AAU16619.	
DR		
PT	New nucleic acid molecules encoding 461 human secreted proteins for	
CC	diagnosing, preventing, treating or ameliorating medical conditions and	
CC	used as food additives or preservatives -	
PS	Claim 1; SEQ ID No 785; 980bp. English.	
XX		
CC	The invention relates to isolated nucleic acid molecules and their	
CC	encoded secreted proteins. The nucleic acids and proteins are used to	
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,	
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They	
CC	are also used in diagnosing a pathological condition or susceptibility	
CC	to a pathological condition. Antibodies to the proteins can also	
CC	be used in alleviating symptoms associated with the disorders and in	
CC	diagnostic immunoassays e.g. radioimmunoassays or enzyme linked	
CC	immunosorbant assays (ELISA). Disorders which are diagnosed or treated	
CC	include autoimmune diseases e.g. rheumatoid arthritis,	
CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,	
CC		

CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Alignment Scores:

Pred. No.:	1,65e-19	Length:	424
Score:	564.00	Matches:	105
Percent Similarity:	85.71%	Conservative:	15
Best Local Similarity:	75.00%	Mismatches:	19
Query Match:	15.50%	Indels:	2
DB:	22	Gaps:	1

US-09-700-590A-22 (1-688) x AAS26606 (1-424)

QY 51 SerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGlnTyrLeuLeuArgValVal 70
DB 3 TCCAAACCCCGAAGCCGCGCCGCAACACCACTGACGATGACGAAATGTGTGTG 62
QY 71 LeuLysThrLeuTyrPheLysGlnPheAlaTyrProPheGlnGlnProValAspAlaVal 90
DB 63 GTGAAGAGCTGTGAAACACAGTTGCTGCTTCTTCAACAGCCCGTGAGCGCATC 122
QY 91 LysLeuAsnLeuProAspTyrTyrLysLysLysLysThrProMetAspMetGlyThrIle 110
DB 123 AAATTGAACCTGCCGATTCATCAATAATTAATAAACCAGATGATATGGGACTATT 182
QY 111 LysLysArgLeuGlnAsnAsnTyrTyrTyrAsnAlaGlnGlnCysIleGlnAspPheAsn 130
DB 183 AAGAGAGACTAGAAATATATTATTATGAGTGCAGACGATGATATGACGACTTCAAC 242
QY 131 ThrMetPheThrAsnCysTyrIleTyrAsnLysProGlnAspAspIleValIleMetAla 150
DB 243 ACCATGTTTCAAAATGTTTCAATTATMACAGCCCAAGATGACATAGTCTAATGGCC 302
QY 151 GlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGlnLeuProThrGlnGlnThrGlu 170
DB 303 CAAGCTTTAGAAATAATTTTCTACAAAAGTGGCCCAAGATGCCCAAGAGAAAGTTGAA 362
QY 171 IleMetIleValGlnAlaLysGlyArgGlyArgGlyArgLysGlnThrGlyThrAlaLys 190
DB 363 ATT---ATTACCCCTGCTCCAA-AGGGGCAAAAGTGTGGAAGCCGCTGGGGAGCCCAAG 418

RESULT 13

AAZ80432

ID AAZ80432 standard; cDNA; 450 BP.

AC AAZ80432;

DT 07-APR-2000 (first entry)

DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:516.

XX Human: gene expression product; diagnosis; tumour; colon cancer;
XX colorectal adenocarcinoma; cell line SW480; cell proliferation;
XX cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
XX hyperplasia; de.

OS Homo sapiens.

XX MO9964576-A2.

XX 16-DEC-1999.

XX

PF 09-JUN-1999; 99WO-1B01062.

XX 10-JUN-1998; 98US-0088801.

XX (FARB) BAYER CORP.

PI Endege WO, Steimann KE, Astle JH, Burgess CC, Bushnell SE;

PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;

PI Schlegel R;

XX MPI; 2000-087220/07.

XX Novel nucleic acids, used to develop products for the diagnosis and

XX treatment of disorders involving unwanted cell proliferation,

XX particularly cancers, especially colon cancer

XX Claim 15; Page 336; 469pp; English.

XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from

XX the human colorectal adenocarcinoma (colon cancer) cell line SW480. The

XX cDNA clones can be used to generate antisense oligonucleotides which

XX can be used for antisense therapy. Methods and products from the present

XX invention can be used for identifying and/or classifying cancerous cells

XX present in a human tumour, particularly in solid tumours, e.g.

XX carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones

XX can be used for developing agents for the diagnosis and treatment of

XX disorders involving unwanted cell proliferation, such as neoplasia,

XX dysplasia or hyperplasia.

XX Sequence 450 BP; 122 A; 127 C; 104 G; 96 T; 1 other;

XX SQ

XX Alignment Scores:

Pred. No.:	1.16e-15	Length:	450
Score:	484.50	Matches:	96
Percent Similarity:	75.00%	Conservative:	12
Best Local Similarity:	66.67%	Mismatches:	27
Query Match:	13.31%	Indels:	9
DB:	21	Gaps:	3

US-09-700-590A-22 (1-688) x AAZ80432 (1-450)

QY 286 LysGlyValLysArgLysAlaAspThrThrThrProThrThrIle----- 300
DB 3 AAAGCGTAAAGCGAAGAGAGATCTACCAACCCCTACACCTACACCTTGGCTCT 62
QY 301 ---AspProIleHisGlnLysProProSerLeuProProGlnProLysThrThrLysLeu--- 318
DB 63 GGTCTCCAGCTAGCCCTCTGGAGTCTT-----GAGCTTAAGGACAGACGCGCTTCCC 116
QY 319 GlyLysArgGlnLysSerSerArgProValLysProProLysLysAspValProAspSer 338
DB 117 CCTATGCGTAGAAGAGTGTGTCGCCCATCATCAAGCCCAAGCAAGATCTGACTCT 176
QY 339 GlnGlnHisProAlaProGlnLysSerSerLysValSerGlnGlnLysCysCysSer 358
DB 177 CAGCAACAACACACAGAGCTCTAAGAAAGAAAGCTTCAAGAACAGTTAAACATGCAAT 236
QY 359 GlyIleLeuLysGlnMetPheAlaLysLysLysAlaLysLysLysLysLysLysLys 378
DB 237 GGCATTTTGAAGAGTACTCTCTTAAGAGCATGCGCTATGCTTGGCTTCTATATAA 296
QY 379 ProValAspValGlnAlaLeuGlnLysLysAspTyrCysAspIleIleLysHisPomet 398
DB 297 CCAATGAGATCTTCTGCACTTGGCTGCATGACTACCAAGACATCTTAAGACACCCCAAG 356
QY 399 AspMetSerThrIleLysSerLysLysGlnAlaArgGlnLysArgAspAlaGlnGlnLys 418
DB 357 GACCTCAGACATGTCAAGCGAAAGATGAGAACCGGATTTACCGGATGACAGAGATT 416
QY 419 GlyAlaAspVal 422
DB 417 GCTGCTGATGTA 428

RESULT 14
ABO60244/c
ID ABO60244 standard; cDNA; 560 BP.
XX
AC ABO60244;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:3939.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (PARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TV, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABO56306 to ABO60787 represent isolated nucleic acids (1) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABO60776 to ABO60787 nucleic acid sequences. (1) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (1) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (1) in a cell. A probe/primer derived
CC from (1) can be used for determining the presence of a nucleic acid which
CC hybridises to (1), and for determining the phenotype of cells in a sample
CC of cells from a patient. (1) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence of
CC state of other type of cancer, in antisense therapy, to generate
CC microarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (1) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 560 BP; 135 A; 121 C; 152 G; 141 T; 11 other;

Alignment Scores:
Pred. No.: 2.99e-14 Length: 560
Score: 457.00 Matches: 92
Percent Similarity: 67.74% Conservative: 13
Best Local Similarity: 59.35% Mismatches: 28
Query Match: 12.56% Indels: 22
DB: 24 Gaps: 3

US-09-700-590A-22 (1-688) x ABO60244 (1-560)

Qy 281 ProVallysrhlylsygllyVallylsarglylsalaspThr-----ThrThrPro 297
Dy 443 CMTTGGCAAGAAAGAAAGCGCTAAAGGAAAGAGATANTTACCACCCCTTACACT 384
Qy 298 ThrThrleasPProIleHisgluPProPser----- 308
Db 383 ACAGCATCTTGCTCTGTTCTCCAGCTACCCCTTGGAGTCTTGAGCNTAAGCA 324

Qy 309 -----LeuProPogluPProlysrThrThrLysLeuGlyGlnArgArgIuSerSerArg 326
Db 323 GCAGCGCTTCCCTC-----ATGCGTAGAGAGAGTGTGCG 288
Qy 327 ProVallysrProPProlylsyAspValProAspSerGlnGlnHisProAlaPProGluLys 346
Db 287 CCCATCAAGCCCGCCAGCAAGACTTGCTGACTCTGACCTAGCAACACCCAGAGCTTTAAG 228
Qy 347 SerSerlysaSerGlnGlnLeuLysCysCysSerGlyIleLeuLysGluMetPheAla 366
Db 227 AAAGAAAGCTTTTCAGAAACGTTAAACATTGCATGCGATTTTGAAGAGATTACTCTCT 168
Qy 367 LysLysHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 386
Db 167 AAGAGAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 108
Qy 387 LeuHisAspPProlysrAspPProlysrAspPProlysrAspPProlysrAspPProlysr 406
Db 107 CTGCATGACTACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 48
Qy 407 LeuGluAlaArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArg 421
Db 47 ATGGAGAACCCGTATTAACCGGATGCAAGAGTTCTCTGAT 3

RESULT 15
AAx29128
ID AAX29128 standard; cDNA; 642 BP.
XX
AC AAX29128;
XX
DT 04-JUN-1999 (first entry)
XX
DE Polynucleotide RTP483 having hypoxia regulated activity.
XX
KW Hypoxia; angiogenesis; apoptosis regulated activity; ischemia; tumour;
KW hypoxic retina; ss.
XX
OS Rattus sp.
XX
PN W09909046-A1.
XX
FD 25-FEB-1999.
XX
PF 21-AUG-1998; 98WO-US17297.
XX
PR 21-AUG-1997; 97US-0056453.
XX
PA (KOHN/) KOHN K I.
PA (QUAR-) QUARK BIOTECH INC.
PI Elinat P, Skalter R;
XX
DR WPI; 1999-180964/15.
XX
PT New polynucleotides with hypoxia, angiogenesis and apoptosis
PT regulated activity - useful for diagnosis of ischemia, and for
PT treatment involving gene therapy

Claim 1; Page 55-56; 72pp; English.

Sequences AAX29126 to AAX29136 represent isolated polynucleotides with hypoxia, angiogenesis and/or apoptosis regulated activity. The invention provides a method for diagnosing ischemia from bodily fluid samples by identifying at least one expressed gene, or identifying at least one up-regulated gene in a tissue sample. In situ analysis was performed for the above sequences in solid tumours and hypoxic retinas. The expression of the genes is activated in tumours around the hypoxic region in the tumour centre, therefore are hypoxia-regulated in vivo. Up regulation promotes angiogenesis required for tumour growth. The genes, proteins, antagonists, antisense oligonucleotides and dominant negative peptides are useful for regulating angiogenesis, apoptosis and hypoxia.

Sequence 642 BP; 169 A; 180 C; 140 G; 151 T; 2 other;

THIS PAGE BLANK (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 18:49:30 ; Search time 148 Seconds
(without alignments)
3262.724 Million cell updates/sec

Title: US-09-700-590A-22

Perfect score: 3639
Sequence: 1 MSASGPGTRLRNLPVMDG.....ALCHLLFAEKETFKLRKLM 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cg2_1/USPTO_spool/US09700596/rnatat_14032003_140631_3291/app.query.fasta_1.839
-DB=Published Applications_NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09700590 @CGN 1.1 80 @rnatat_14032003_140631_3291
-NCPU=6 -ICPU=3 -NO.XLPEXY -NO.NMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	51.9	4664	10	US-09-880-107-1657 Sequence 1657, App
2	688	18.9	746	10	US-09-764-864-372 Sequence 372, App
3	564	15.5	424	10	US-09-764-864-785 Sequence 785, App
4	484.5	13.3	450	10	US-09-879-536-516 Sequence 516, App

5	362.5	10.0	429	10	US-09-876-889-243 Sequence 243, App
6	323.5	8.9	7326	12	US-10-109-886-7 Sequence 7, App1
7	308.5	8.5	547	10	US-09-815-343-213 Sequence 213, App
8	308	8.5	8147	12	US-10-109-886-9 Sequence 9, App1
9	272.5	7.5	358	10	US-09-783-590-1966 Sequence 1966, App
10	262.5	7.2	1855	10	US-09-864-761-19708 Sequence 19708, App
11	262.5	7.2	1958	10	US-09-864-761-2927 Sequence 2927, App
12	236	6.5	5037	9	US-09-808-880-1 Sequence 1, App1
13	226	6.2	1140	9	US-09-938-842A-1265 Sequence 1265, App
14	226	6.2	1140	9	US-09-887-576-652 Sequence 652, App
15	225	6.2	2653	10	US-09-764-864-373 Sequence 373, App
16	221.5	6.1	2825	10	US-09-822-889A-258 Sequence 258, App
17	221	6.1	2832	10	US-09-764-864-371 Sequence 371, App
18	220	6.0	2138	10	US-09-764-864-784 Sequence 784, App
19	220	6.0	2283	9	US-09-938-842A-1691 Sequence 1691, App
20	220	6.0	1989	9	US-09-887-576-645 Sequence 645, App
21	218.5	6.0	1998	9	US-09-894-844-6 Sequence 6, App1
22	218.5	6.0	2050	10	US-09-791-171-69 Sequence 69, App1
23	217.5	6.0	2824	12	US-10-124-557-13 Sequence 13, App1
24	217.5	6.0	3066	12	US-10-124-557-83 Sequence 83, App1
25	217.5	6.0	3117	12	US-10-124-557-73 Sequence 73, App1
26	217.5	6.0	3148	12	US-10-124-557-57 Sequence 57, App1
27	217.5	6.0	3420	12	US-10-124-557-103 Sequence 103, App
28	217.5	6.0	3813	12	US-10-124-557-43 Sequence 43, App1
29	217.5	6.0	3936	12	US-10-124-557-41 Sequence 41, App1
30	217.5	6.0	3942	12	US-10-124-557-141 Sequence 141, App
31	217.5	6.0	3945	12	US-10-124-557-49 Sequence 49, App1
32	217.5	6.0	3963	12	US-10-124-557-45 Sequence 45, App1
33	217.5	6.0	3963	12	US-10-124-557-59 Sequence 59, App1
34	217.5	6.0	4065	12	US-10-124-557-47 Sequence 47, App1
35	217.5	6.0	4086	12	US-10-124-557-39 Sequence 39, App1
36	217.5	6.0	4092	12	US-10-124-557-51 Sequence 51, App1
37	217.5	6.0	4215	12	US-10-124-557-61 Sequence 61, App1
38	217.5	6.0	4575	12	US-10-044-090-303 Sequence 303, App
39	217.5	6.0	5008	12	US-10-124-557-1 Sequence 1, App1
40	217.5	6.0	15720	9	US-10-025-380-1058 Sequence 1058, App
41	217.5	6.0	15720	10	US-09-922-217-1058 Sequence 1058, App
42	217.5	6.0	15720	10	US-09-833-263-1058 Sequence 1058, App
43	215.5	5.9	2272	9	US-10-174-590-345 Sequence 345, App
44	215.5	5.9	2272	9	US-10-176-758-345 Sequence 345, App
45	215.5	5.9	2272	9	US-10-175-737-345 Sequence 345, App

ALIGNMENTS

RESULT 1
US-09-880-107-1657
; Sequence 1657, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1657
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D42040
US-09-880-107-1657

Alignment Scores:

Pred. No.: 2.05e-104 Length: 4664
 Score: 1888.00 Matches: 404
 Percent Similarity: 67.08% Conservative: 81
 Best Local Similarity: 55.88% Mismatches: 142
 Query Match: 51.88% Indels: 97
 DB: 10 Gaps: 20

US-09-700-590a-22 (1-688) x US-09-880-107-1657 (1-4664)

QY 5 SerGIProGIYThrArgLeuArgAsnLeuProValMetGIAspGILeuGIuThSer 24
 DB 1780 GCAGCACCAGGAGAAAGATTGAAACCCCTCTCTTGTATGAGGCGTTGAGAGCCCC 1839
 QY 25 GlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThrAsn 44
 DB 1840 ACATGGGCTGGGCTGCT---TTCACACTTACCCCTGCC-----AAC 1881
 QY 45 ProProProProGIuThrSerAsnProAsnLySProLySAlaGlnThrAsnGlnLeuGln 64
 DB 1882 CCACCACCCCGGAGGTGTCACATCCCAAAAGCCAGAGACGATTACCCACAGCTGCAG 1941
 QY 65 TyrLeuLeuArgValIValLeuLeuSerThrLeuTrpLysHISGlnPheAlaTrpProPheGln 84
 DB 1942 TACCTTACCAAGGATGATGATAAGGCTGTGGAACATCAGTTGCGATGGCCATTCGCG 2001
 QY 85 GlnProValAspAlaValIValLeuAsnLeuProAspTyrTyrIleLeuLeuTherPro 104
 DB 2002 CAGCTGTGATGCTGCTCAAACTGGGCTACCGGATTCACAAATATATAAGGAGCCT 2061
 QY 105 MetAspMetGIYThrIleValLeuValArgLeuGluAsnAsnTyrTyrTrpAsnAlaGlnIu 124
 DB 2062 ATGACATGTGGTACTATTAAGAGGAGACTTCAAAACATTTATTTGGCTGCTTCAGAG 2121
 QY 125 CysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsnLySProGIYAsp 144
 DB 2122 TGTATGCAAGATTTTATATACATGTTTCACAACTGTTACATTTCACACAGCCACTGAT 2181
 QY 145 AspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLysIleAsnGluLeu 164
 DB 2182 GATTTGTGCTTAATGGGCAACAAAGCTGGAAAGATATTCCTTACAGAAAGTTGCATCATG 2241
 QY 165 ProThrGluGIuThrGluIleMetIleValGlnAlaLysGIY-----ArgGIY-- 180
 DB 2242 CCACAGAGAAACAAAGAGCTGTAGTACCATCCCTTAAGAACAGCACAGAGAGGGGCC 2301
 QY 181 -----ArgGIYArgLysGIuThrGluThrAlaLysProGIYValSerThr 195
 DB 2302 AAGTTGGCAGCGCTCCAGGGCAAGTGTTCACAGTCCCATCAGGTGCTGCGCTCTTCT 2361
 QY 196 ValProAsnThrThrGlnAlaSerThrProProGIuThrGlnThrProGlnProAsnPro 215
 DB 2362 GTGTCACACAGAGCCCTGTATACCTCTCCACCTGAGATA-----CCT 2403
 QY 216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235
 DB 2404 ACCATGTCTTCACATATCCCAACCA-----TCAGTCAATTTCC 2442
 QY 236 ThrProValMetThrValValProProGlnProLeuGlnThrProProValProPro 255
 DB 2443 TCTCCACTTCTCAAGTCTTGAC-----TCTGCTGACCCCGCTCTTCT 2490
 QY 256 GlnProGlnProProAlaProAlaProGlnProValGlnSerHisProProIleIle 275
 DB 2491 GTTACTGAGAGCTCTCCAGCC----- 2511
 QY 276 AlaAlaThrProGlnProValLysThrLysGIYValIValArgLysAlaAspThrThr 295
 DB 2512 -----CAGCCCTTGCCCAAGAAAGCGGTAAAGCGGAAAGCAGATACTACC 2559
 QY 296 ThrProThrThrIle-----AspProIleHisGluProProSerIeu 309
 DB 2560 ACCCTTACCTACAGCAGCATTTGGCTCTGTTCTCCAGCTAGCCCTCTCGGAGTCTT 2619

QY 310 ProProGluProPolysThrThrLysLeu--GlyGlnArgArgGluSerSerArgProVal 328
 DB 2620 -----GAGCTTAAGGACAGCAGCGCTTCCCTTATGATGAGAGAGTGTGGCCATC 2673
 QY 329 LysProProLysAspValProAspSerGlnGlnHisProAlaProGIuLysSerSer 348
 DB 2674 AAGCCCCAGCAGAAAGCTTGCTGACTTCCAGCAACACAGAGCTCTTAAGAAAGCA 2733
 QY 349 LysValSerGIuGlnLeuLysCysCysSerGIYIleLeuLysGIuMetPheAlaLysLys 368
 DB 2734 AAGCTTCAGAAAGCTTAATAACATTGCAATGGCATTTTGAAGGAGTATCTTAAGAG 2793
 QY 369 HisAlaAlaTyrAlaTrpProPheTyrLysProValAspValGluAlaLeuGIYLeuHis 388
 DB 2794 CATGCTGCCTTATGCTTGCTTATTAACCAAGTGAATCTTCTGCACTTGGCTTCAT 2853
 QY 389 AspTyrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGIu 408
 DB 2854 GACTTACATGACATCATTAAGCACCCTGAGACTTCAGCACTGTCAAGCGAGATGGAG 2913
 QY 409 AlaArgLysTyrArgAspAlaGlnGluPheGIYAlaAspValArgLeuMetPheSerAsn 428
 DB 2914 AACCTGATTAACGGGATGACAGAGACTTGGCTGATGTAGCGCTTATGTTCTCCAC 2973
 QY 429 CysTyrLysTyrAsnProProAspHisGluValAlaMetAlaArgLysLeuGlnAsp 448
 DB 2974 TGCTATAGTACATACCCCAAGATCAGAGATGCTGTGCAATGGCAGCAAAAGCTACAGAT 3033
 QY 449 ValPheGluMetArgPheAlaLysMetProAspGIuProGIuGIuProValIValAlaVal 468
 DB 3034 GTATTTGAAGTCCGTTATGCGAAGATGCCAGATGACACACTAGAACAGAGGCGCTTACCA 3093
 QY 469 SerSerProAlaValProProPro---ThrLysValValAlaProProSerSerSerAsp 487
 DB 3094 GTCTTACTGCGACGCCCTCGCTGGCCCAATCGCTTCAGAGTCTCCAGTGAAGAA 3153
 QY 488 SerSerSerAspSerSerSer-----Asp 495
 DB 3154 AGTAGCAGTGAAGCTCTCTCTGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAA 3213
 QY 496 SerAspSerSerThrAspAspSerGIuGluGluArgAlaGlnArgLeuAlaGluLeuGln 515
 DB 3214 GAAGAGAGTGAAGCTCAGACTCAGAGAGAAAGAGGCTCATCTTATGAGAACTACAG 3273
 QY 516 GlnGlnLeuLysAlaValHisGluGlnLeuAlaLeuSerGlnProGlnGlnAsnLys 535
 DB 3274 GAACAGCTTGGCGCAGTACATGAACAACTGCTCTGCTCCAGGGTCCAAATATCCAAAG 3333
 QY 536 ProLysLysLysGIuLysAspLysLysGIuLysLysLys-----GluLysHis 551
 DB 3334 CCCAAGAGGAAA---AGAGAGAAAAAGAGAAAAAGAAACGAGAGCAGAGAACAT 3390
 QY 552 LysArgLysGIuGluValGluGluAsnLysLysSer---LysAlaLysGIuProPro--- 569
 DB 3391 CGAGCGGAGCTGGGGCCGAGTGAAGATGACAGGGGCTTAGGGCCGCCCACTTCAA 3450
 QY 570 ProLysLysThrLysLysAsnAsnSerSerAsnValSerLysLysGIuProAla 589
 DB 3451 CCTAAGAAAGTCAAGAAAGCAAGTGCAGTGGGCTGCGCTTAAAGGCGCTTCT 3510
 QY 590 -----PrometLysSerLys 594
 DB 3511 GCGTTTGACCTTCTGAGAGAAAGTGCACCAAGCTCCCAAAAAGGCACAAAGACAGCC 3570
 QY 595 ProPro-----ProThr---TyrGluSerGIuGluLysAspLysCysLysProMetSer 611
 DB 3571 CCACCTGCTCCTTACAGGTTATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3630
 QY 612 TyrGIuGluLysArgGlnLeuSerLeuAspIleAsnLysLeuProGIYGIYLysLeuGIY 631
 DB 3631 TACATGAGAGCGGACCTGAGCTGAGCATCAACAATTAATCTGGGAGAGAGCTGGCC 3690
 QY 632 ArgValValHisIleIleGlnSerArgLysProSerLeuLysAsnSerAsnProAspGIu 651

```
|||||
Db 3691 CGAGTGTGATTAATCCAGGAGGAGCCCTCTTACGTATCAACCCAGAGAG 3750
QY 652 ILEGLIILASPPHEGLINRILEULYSPROSETHLEULRGILUENGLIYALALEUCYS 671
Db 3751 ATTGAATTTGATTTTGAACACTCAAGCCATCCACACTTAGAGAGCTTGA-GGGCTAATG 3809
QY 672 HLEULEU 674
Db 3810 CCTTCTCG 3818

RESULT 2
US-09-764-864-372
; Sequence 372, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 372
; LENGTH: 746
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (387)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (646)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (665)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (736)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-372

Alignment Scores:
Pred. No.: 1,5e-33 Length: 746
Score: 688.00 Matches: 144
Percent Similarity: 66.40% Conservative: 22
Best Local Similarity: 57.60% Mismatches: 61
Query Match: 18.91% Indels: 24
DB: 10 Gaps: 6

US-09-700-590A-22 (1-688) x US-09-764-864-372 (1-746)
QY 51 SerAsnProAsnLysProLysArgInThrAsnGlnLeuGlnTyLeuLeuArgValVal 70
Db 3 TCCAAACCCCAAGAGCCGCGCCGCAAGACCAACAGTCAGTGCAGTAATGTGCTG 62
QY 71 LeuLysThrLeuTrpLysHisGlnPheAlaTrpProPheGlnInProValAspAlaVal 90
Db 63 GTCAAGACGCTCGGAACACACAGTTCGCTGCGCTTCAACAGCCCGTGGACGCAATC 122
QY 91 LysLeuAsnLeuProAspTyTyTyLysIleIleLysThrProMetAspMetGlyThrIle 110
Db 123 AAATTGAACCTGCGGATTAATCAATAATTAATAAACCAATGATATGGGACTATT 182
QY 111 LysLysArgLeuGlnLysAsnEntyTyTyTrpAsnAlaGlnGlnLysIleGlnAspPheAsn 130
Db 183 AAGAAGAGACTAGAAAATATATTATTTGAGTGCAGAGCAATGATATGAGACTTCAAC 242
QY 131 ThMetPheThrAsnCyTyTyTyTyAsnLysProGlyAspAspIleValIleMetAla 150
Db 243 ACCATGTTTACAAATGTTTACATTATTAACAAGCCACAGATGACATAGCTAATGGCC 302
```

```
QY 151 GluAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGlnLeuProThrGlnGlnThrGln 170
Db 303 CAAGCTTTAGAGAAAATTTTCTACAAAAAGTGCCCAAGATGCCCAAGAGAAATTGAA 362
QY 171 IleMetIleValGlnAlaLysGlyArgGlyArgGlyGlnThrGlyThrAlaLys 190
Db 363 ATTATTACCCCTGCTCCAAAGGNCANAGTCCAGAGCCGGT-GGGAGAGCCAGAGC 421
QY 191 ProGlyValSerThrValProAsnThrThrGlnAlaSerThrProGlnInThrGlnThr 210
Db 422 GCAGTACACAGAGAGTGGCGGCGGTCTCTGCTCC-----CCAGGAGCC 469
QY 211 Pro---GlnProAsnProProProValGlnAlaThrProHisProPheProAlaValThr 229
Db 470 CCTTTCAAGAGCGTGCCCGCCACCGCTCCAGACGCC----- 508
QY 230 ProAspLeuIleValGlnThrProValMetThrValValProGlnInProLeuGlnThr 249
Db 509 -----GTCACTGCTGCCACCCCTGTACCAACCATCACTGCAAAC-----GTCACG 553
QY 250 ProProProValProProGlnProGlnProProProAlaProAlaProGlnProValGln 269
Db 554 TCGGTCCCAAGTCCCGCCCACTGCGGCCCACTCTCT-----CCTGCCACA 598
QY 270 SerHisProProIleIleAlaAlaThrProGlnProValLysThrLysGlyValLys 289
Db 599 CCCATGCTCCCGGTGGTCCCTCCACGCCA---CCTGTGTCAGAAAGGCGTTGAA 655
QY 290 ArgLysAlaAspThrThrThrThrProThrThr 299
Db 656 CGGAAGACANACACACACCTCCAGACT 685

RESULT 3
US-09-764-864-785
; Sequence 785, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 785
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-785

Alignment Scores:
Pred. No.: 2,15e-26 Length: 424
Score: 564.00 Matches: 105
Percent Similarity: 85.71% Conservative: 15
Best Local Similarity: 75.00% Mismatches: 19
Query Match: 15.50% Indels: 2
DB: 10 Gaps: 1

US-09-700-590A-22 (1-688) x US-09-764-864-785 (1-424)
QY 51 SerAsnProAsnLysProLysArgInThrAsnGlnLeuGlnTyLeuLeuArgValVal 70
Db 3 TCCAAACCCCAAGAGCCGCGCCGCAAGACCAACAGTCAGTGCAGTAATGTGCTG 62
QY 71 LeuLysThrLeuTrpLysHisGlnPheAlaTrpProPheGlnInProValAspAlaVal 90
Db 63 GTCAAGACGCTCGGAACACACAGTTCGCTGCGCTTCAACAGCCCGTGGACGCAATC 122
QY 91 LysLeuAsnLeuProAspTyTyTyTyLysIleIleLysThrProMetAspMetGlyThrIle 110
Db 123 AAATTGAACCTGCGGATTAATCAATAATTAATAAACCAATGATATGGGACTATT 182
```

```

Oy      111 LysLysrArGlauGluAuaAsnTyTfTrpAsnIacInglucYslIegInAspPheAsn 130
Db      183 AAGAAGACACTGAAGAAATAATTAATTATTTGGAGTCACAAGCAAGATATGACAGACTTCAC 242
Oy      131 ThirnetPherthaenCySTyTlIEtYAsnlSpFroJlYaSPaPIleValIleumetaJa 150
Db      243 ACCATGTTTAACAAATTGTTACATTATTTATACAAAGCCCAACAGATGACATATGCTAATGCC 302
Oy      151 GluAlaleuglulvSleupHeuleGlnlyslleasnGlueupProthrnglulgutThrngu 170
Db      303 CAAGCTTTAGAGAAAATTCTTCACAAAAAATGTGCACAGATGTCCCCCAGAGGAAGTTGAA 362
Oy      171 IleMetIleValIGlnAlaIeYsgLYArGclYArGclYArGlsuThrgluThyrAlaIaYs 190
Db      363 ATT--ATTACCCTCGCTCCAA-AGGGGCAAAGGTGGAAGCCGGCTGGCGGAGGCCAG 418

RESULT 4
US-09-879-536-516
Sequence 516, Application US/09879536
Patent No. US20020144298A1
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/879,536
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/088,801
PRIOR FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 516
LENGTH: 450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(450)
OTHER INFORMATION: n = A,T,C or G
US-09-879-536-516

Alignment Scores:
Pred. No.: 1,28e-21 Length: 450
Score: 484.50 Matches: 96
Percent Similarity: 75.00% Conservative: 12
Best Local Similarity: 66.67% Mismatches: 27
Query Match: 13.31% Indels: 9
Gaps: 3
DB: 10

US-09-700-590A-22 (1-688) x US-09-879-536-516 (1-450)
Oy      286 LysgLyVallyArGLyArGLaAlaSPrThrThrThrProThrHrile----- 300
Db      3 AAAAGCGTAAAGCGGAAAGACAGATCTACACACCCTTACACTACAGCATCTTGCTCCT 62
Oy      301 ---AspPrIoleHisGluProProSerleuProProGluProLySthThrLYSLeu--- 318
Db      63 GGTTCTCCAGCTAGCCCTCCTGGAGTCTT-----GAGCTTAAGGAGCAGCGCTTCCC 116
Oy      319 GlyGlnArGaRGrGluSeSerHrpProValysPrrProLySlYsAspAlProAsSer 338
Db      117 CCTATGCGGTAGAGAGTGTGTGCGCCCAATCAAGCCCAAGCAAGACTTGCTGACTCT 176

```

Oy	339	GnGlnHhSProAlaP	roGluLySseSerLyVaLSe	rGluGluInLeuLyS	CysCySse	336
Db	177	CAGACACACAC	CAGACCTTAA	GAAGAAAGAAAG	ATTTCACAAACGTTTAA	ACATTGCAAT 236
Oy	359	GlyIleLeuLyS	GumetPheAlaLyS	hISaLaAlaLy	rAlaTrpProPheTyLyS	378
Db	237	GGCATTTTGA	AGAGCTTAA	CTCACTCA	TAACACATGCTGCTTGGCTTTCTAT	TAAA 296
Oy	379	ProValaSP	ValGluAlaLeuGlyLeu	hISAsPTrCyS	ApIleIleLyS	hISpPromet 398
Db	297	CCAGTGGATG	CTTCTGCACCTTG	CTGCATAC	CTACATCACTATTAA	GACACCCATG 356
Oy	399	AspMetSerTrIle	LySseSerLyS	LeuGluAlaArgGly	TrpArgSPAlaGInGluPhe	418
Db	357	GACCTCAG	CACTGTCAAGCGG	AGATGGAACCG	TGATTACCGGAGTACAGAGACTTT	416
Oy	419	GlyAlaSP	AlVal 422			
Db	417	GCTGCTGAT	GTGA 428			
RESULT 5						
US-09-876-889-243						
; Sequence 243, Application US/09876889						
; Patent No. US20020076715A1						
GENERAL INFORMATION:						
; APPLICANT: Benson, Darin R.						
; APPLICANT: Lodes, Michael J.						
; APPLICANT: Mitcham, Jennifer L.						
; APPLICANT: King, Gordon E.						
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN						
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS						
; FILE REFERENCE: 210121.466C3						
; CURRENT APPLICATION NUMBER: US/09/876,889						
; CURRENT FILING DATE: 2001-06-06						
; NUMBER OF SEQ ID NOS: 353						
; SOFTWARE: FastSeq for Windows Version 3.0						
; SEQ ID NO 243						
; LENGTH: 429						
; TYPE: DNA						
; ORGANISM: Homo sapien						
US-09-876-889-243						
Alignment Scores:						
Pred. No.: 2.36e-14 Length: 429						
Score: 362.50 Matches: 79						
Percent Similarity: 68.53% Conservative: 19						
Best Local Similarity: 55.24% Mismatches: 24						
Query Match: 9.96% Indels: 21						
DB: 10 Gaps: 5						
US-09-700-590A-22 (1-688) x US-09-876-889-243 (1-429)						
Oy	538	LySLySGluLyS	AsPlyLySGluLySLyS	hISLySGluLyS	rAlaTrpProPheTyLyS	557
Db	2	AAAAAGGAAAA	---AGAGAAACGG	AGAGCAGAGAAACATCTGGCCCAATTGGGATC	58	
Oy	558	GluGluAsnLyS	LySse---LyS	AlaLySGluProPro--ProLyS	LyS	575
Db	59	GATGAAGTAT	GAATTAAGGGGCGCT	AGAGGCACTGCGCCACCTAGCCCAAGAAATCTTAAGAAA	118	
Oy	576	AsnAsnSerSer	AsnSerAsnVal	-----		583
Db	119	GCAGGTGTGTGGGGT	AGCAATGCTACT	ATACACTCAGCACCCTCGCTTTGGGACTTCCGA	178	
Oy	584	-----SerLyS	LySGluProAlaPro	MetLyS	SerLySProProProThr-----	598
Db	179	GGAACTAC	CAACAAGCTTAA			

```

Qy 618 LeuSerLeuAsp11LeuAsnLysLeuProGlyGlyLysLeuGlyArgValValHisIleIle 637
Db 299 TTAAGCTTGATATCATATAAGTTACTCTGGGAAAGCTGGGTCCAGTAGTACATATCATC 358
Qy 638 GluSerArgGluProSerLeuLysAsnSerAspProAspGluIleGluIleAspPheGlu 657
Db 359 CAAAGCCAGGAAACCTCTCTACGATTCAATCCAGAGAAATTGAGATTGATTGAA 418
Qy 658 ThrLeuLys 660
Db 419 ACACTCAAG 427

RESULT 6
US-10-109-886-7
; Sequence 7, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; TITLE OF INVENTION: ANTAGONIST TO PPAR
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 7326
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(7326)
; OTHER INFORMATION: n at position 1131 is unknown.
US-10-109-886-7

Alignment Scores:
Pred. No.: 9.04e-11 Length: 7326
Score: 323.50 Matches: 135
Percent Similarity: 35.88% Conservative: 72
Best Local Similarity: 23.40% Mismatches: 177
Query Match: 8.89% Indels: 193
DB: 12 Gaps: 27

US-09-700-590a-22 (1-688) x US-10-109-886-7 (1-7326)
Qy 9 ThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGluThrSerGlnMetSerThr 28
Db 2008 ACAAGCTTACATAGCAAGAGCATCTCGGTAACTC----- 2040
Qy 29 ThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaIleSerThrAsnProProPro 47
Db 2041 -----CAGCCAGCTTTACCAAGCTTCTGGGGCTCAAGCCCTGTG 2079
Qy 48 ---ProGluThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGlnTyrLeu 66
Db 2080 ATTCCACCAAGCCAGCTGTGAAGACTCCCAAT----- 2112
Qy 67 LeuArgValValLeuLysThrLeuTyrLysHisGlnPheAlaTyrProPheGlnGlnPro 86
Db 2113 -----GGGCCCTGCTTGTGCA 2130
Qy 87 ValAspAlaValLysLeuAsn-----LeuProAspTyrTyrLysIleIleLysThrPro 104
Db 2131 GTGAATGATGATGAGGTTTCTCAAGAGATGATGATTCATTT-----AACCCA 2175

```

```

Qy 105 MetAspMetCylThrIleLys-----LysArg 113
Db 2176 ATGTCCCTGGAAGAACTTCCAGTCCACAGCAAGCCATGAGACCTGTGACGCTCCCT 2235
Qy 114 LeuGlnAsnAsnTyrTyrTyrPheAlaGlnGlnLysIle----- 126
Db 2236 ATGAACCACTCTGACAGATGAACAGATGAGCTCAAGTTCCTGGGTATGGCATTTCTCCT 2295
Qy 127 -----GlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsnLysPro 142
Db 2296 TCACGATGCTCAGCGCTCCAAATATGATGGGCATCATGCC-----AACAC-- 2443
Qy 143 GlyAspAsp11ValLeuMetAlaGlnAla---LeuGlnLysLeuPheLeuGlnLysIle 161
Db 2344 -----ATTATGGCCAGGACCTACTACAGAACGATTTCTGCCACAG-- 2385
Qy 162 AsnGluLeuProThrGluGluThrGluIleMetIleValGlnAlaLysGlyArg 181
Db 2386 AACCAAGTTTCATCA---TCCAGTGGGGCAATGATGTGAACAGGTGTGGCATGGGCA 2442
Qy 182 GlyArgLysGluThrGlyThrAla-----LysProGlyValSer----- 194
Db 2443 CCAAGCAAGCCAGGCAAGTGTTCACAGAGGTCAAGAACTGGAGCTGTCTCCCTAACCT 2502
Qy 195 -----ThrValProAsnThrThrGlnAlaSerThrProGlnThrGlnThrPro 211
Db 2503 CTGAACATGCTGACCCCAAGCCAGCCAGCTGCTTGCACCAAGTGAACAGTCAACCA 2562
Qy 212 ---GlnProAsnProProPro-----ValGlnAlaThrProHisPro 224
Db 2563 TTGCACCCGACCTCAGCTCTCTGTTCCACAGCTGCTGAGTGGCTCTCTCCAACTCA 2622
Qy 225 ---PheProAlaValThrProAspLeuIleValGlnThrProValMetThrValPro 243
Db 2623 ACGGCACCAAGATGACCT-----CTCAGCCAGCAGCTCCCACTCAAGCA 2670
Qy 244 ProGlnProLeu-----GlnThrProProProValPro----- 254
Db 2671 TCTACTCTGTGTCTCATCTGGGACAGACTCTCAACCACTCTGCTCAGTCCAGCGCT 2730
Qy 255 -----ProGln 256
Db 2731 GCCCAACACAGATACCCCTACAGTCCAGCAGCAGCAGCAGCTGAGTCACTCAAG 2790
Qy 257 ProGlnProProProAlaPro-----AlaPro 265
Db 2791 CCTCAAGACCCCACTGACGACCATCTGTGGCTACTCTCAGTCACTCAAGCAGCAACCA 2850
Qy 266 GlnProValGlnSerHisProProIleIle----- 275
Db 2851 ACGCTGTGATCTCAGCCACTGCGCACACGCGCTTCTCAGCAGCAGCAGCATTTGAT 2910
Qy 276 -----AlaAlaThrProGlnProValLysThrLysLysGlyValLysArgLysAlaAsp 293
Db 2911 AATAGATGCTCTACTCTCTCACTGACAGTGTGAACCAAGTTCACAGCAGCAGCA 2970
Qy 294 ThrThrThrPro-----ThrThrLeuAspProIleHisGlu 305
Db 2971 CCCGATGTGCCATGCTGTGAATGAAGACAGAGCTCAAGCATATGCTGAGCTGAA 3030
Qy 306 ProProSerLeuProProGluProLysThrThrLysLeuGlyGlnArgArgLysSer 325
Db 3031 CTACAGTAATCCAAAGGGGGAACCTCGTGTGAAGATGAAGAGATTTCACAGGTTCT 3090
Qy 326 ArgProValLys-----ProProLysLysAspValProAspSerGlnGln 340
Db 3091 TCCCAAGTAAAGAAAGACAGATACAGACAGACAGAAAGTCAAGCAATGAGATGAA 3150
Qy 341 HisProAlaProGlu-----LysSerSerLys 349
Db 3151 GAAAGAAACCTGAAATGAAGTGAAGCTAAAGAGAAAGAAAGAAAGTTCGAAAGCAG 3210
Qy 350 ValSerGluGlnLeuLysCysCysSerGlyIleLeuLysGluMetPheAlaLysLys--- 368

```


QY 197 ProAsnThrGlnAlaSerThrProGlnThrGlnThrPro---GlnProAsnPro 215
Db 3330 CCTCAGGCCGCGCTACCTGCTCCAGTGCAGACGACCTGCACCAACGACC 3389
QY 216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235
Db 3390 CCTCTGCTCCACGGCT-----GCTGGCATGCTCTCCAGCACACG 3434
QY 236 ThrProValMetThrValValProGlnPro---LeuGlnThrProProValPro 254
Db 3435 ACACCA---CCTGGATGATGATCTCTCCACGACGACCTCCCACTACCACTACTCT 3491
QY 255 -----ProGlnProGlnProProProAlaProAlaProGlnPro 267
Db 3492 GTGTGCTCTCCGCGACGACCTCCACCCGACCTGCTCAGTGCCTCAGTACCCAA 3551
QY 268 ValGlnSerHisProProIleIleIleAlaIle-----ThrProGln----- 280
Db 3552 ACCCAGAGCACCCTCAGTCAGCGACGACGCCAGGCCAGGTGACCCGACCTCAA 3611
QY 281 ---ProValThrLysLysGlyValLysArgLysAlaAspThrThrThrProThrThr 299
Db 3612 ACCCAGATTAGCCCCGCTGTGCTGCTACCTCAGTCATCGACGACGACCGACG--- 3668
QY 300 IleAspProIleHis---GlnProProSerLeuProProGlnProLysThrLysLeu 318
Db 3669 -----CCTGTGACGCGCCGACCTCTGCGACACCGCTTCCAGCAGACGACGAT 3722
QY 319 GlyLysArgLysSerSerArgProVal-----LysProPro 331
Db 3723 GATTAACGAGTCCCTACCCCTCTCGGTGCGACGCGAAMCAATCCCGACGACCA 3782
QY 332 LysLysAspValPro-----AspSerGlnGln 340
Db 3783 GGACTTACCTACTCTGTCTGGAATGAATGAACGAGACCCAGAGAGGACACTGAGCC 3842
QY 341 HisProAlaProGlnLys-----SerSerLysValSerGlnLysLysCys 356
Db 3843 GATCTGTGTAATCTCAAGGGAGGCCAGCGTCTGATGATGAGAGGATTCGAAAGA 3902
QY 357 CysSerGlyIleLeuLysGlu----- 363
Db 3903 GCTTCCCAAGTTAAGAAAGAACAGACATAGCAGACGAGAATCAGAACATGGAAGTG 3962
QY 363 ----- 363
Db 3963 GATGAAAAAACTCTGAAGTGAAGTGAAGTGAAGAGAGAGAGTAGCATTAAC 4022
QY 363 ----- 363
Db 4023 GGCACAGCCTCTCAGTACATCTCTTCGACGCGGCAAAAAATCTTTAAACAGAG 4082
QY 364 -----MetPheAlaLysLysHisAlaIleTyr-----Ala 373
Db 4083 GAGTTACGCGAGCCCTCATGACCAACCTTAGAGCACTGATTCACAGACGACGAGTCA 4142
QY 374 TyrProPheTyrLysProValAspValGlnAlaLeuGlnLysHisAspTyrCysAspIle 393
Db 4143 TTAACCTTTCGCGAGCCTGATGATCCGAGCTCTCGAATCTCCAGATTTTGAATC 4202
QY 394 IleLysHisProMetAspMetSerThrIleLysSerLysLeuGlnAlaArgGlnTyrArg 413
Db 4203 GTAAGATATCCATGAGCCTCTCCACATCAAGCGAAGCTGGACACAGGCGCATTAACA 4262
QY 414 AspAlaGlnGlnPheGlnAlaAspValArgLeuMetPheSerAsnCysTyrLysTyrAsn 433
Db 4263 GAGCCTGCGAGTACGAGCAGACGCTGCTGCTCAATGCTGCTCTATTAAT 4322
QY 434 ProProAspHisGlnValValAlaMetAlaArgLysLeuGlnAspValPheGln 451
Db 4323 CGCAAGACATCCGAGTCTATTAAGTTTGCAGTAAGCTTGACAGAGCTTTGAG 4376
RESULT 9

US-09-783-590-1966
Sequence 1966, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillion, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16 2C1
CURRENT APPLICATION NUMBER: US/09/783,590
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1966
LENGTH: 358
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (41)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (102)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (121)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (158)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (180)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (186)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (203)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (225)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (237)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (329)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (331)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (334)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (336)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature


```

FEATURE:
; OTHER INFORMATION: MAP TO AC005529.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
US-09-864-761-2927

Alignment Scores:
Pred. No.: 1.04e-07 Length: 1958
Score: 262.50 Matches: 126
Percent Similarity: 38.61% Conservative: 96
Best Local Similarity: 21.91% Mismatches: 268
Query Match: 7.21% Indels: 85
DB: Gaps: 20

US-09-700-590a-22 (1-688) x US-09-864-761-2927 (1-1958)
Qy 138 ILETyRAsnLysProGlyAspAspIleValIleuMetAlaGluAlaLeuGluLysLeuPhe 157
Dy 272 GTGTCGAGGAAGTCTGAGAAGAACTGTGATTCGAGGAACAGACAGAGAGG----- 325
Qy 158 LeuGluLysIleAsnGluLeuProThrGluGluThrGluIleMetIleValGluAlaLys 177
Dy 326 --ACCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 382
Qy 178 GLyRAsnGlyAsnGlyAsnGlyAsnGlyAsnGlyAsnGlyAsnGlyAsnGlyAsnGlyAsn 197
Dy 383 GGCAGAGGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 427
Qy 198 AsnThrThrGlnAlaSerThrProGlnThrGlnThrProGlnPro----- 213
Dy 428 ---GAAGAAAGAAAGAAAGTCTCCCGAGAGAGAGGAGGAGTCCCGAGAGAGAGAGCC 484
Qy 214 AsnProProProValGluAlaThrProHisProPheProAlaValThrProAspLeuIle 233
Dy 485 AAGTCAACAGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
Qy 234 ValGlnThrProValMetThrValValProProGlnProLeuGlnThrProProProVal 253
Dy 545 GAAGCAAAATCCCGAGCCGAGAGTCAAGTCCCTGAGAAAGCCCAAGTCTCCAGCAAAAGAA 604
Qy 254 ProProGlnProGlnProProProAlaProAlaProAlaProAlaProAlaProAlaPro 273
Dy 605 GAGGCAAAAG---TCAACCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
Qy 274 IleIleAlaIleThrProGlnProValLysThr-----LysLysGlyValLysArgLys 291
Dy 662 GCTGAGTCAAGTCTCCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 721
Qy 292 AlaAspThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 310
Dy 722 GCTGAGGAGCAAGTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781
Qy 311 ProGlnProLysThrThrLysLeuGlnArgArgGluSerSerArgProValLysPro 330
Dy 782 GCTGAGGAGCAAGTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
Qy 331 ProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSerLysVal 350
Dy 842 GAAAAGGAGCAAGTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
Qy 351 SerGluGlnLeuLysCysSerSerGlyIleLeuLysGluMetPheAlaLysLysHisAla 370
Dy 902 GAG-----AAGGAGAGAGGCC 916
Qy 371 AlaTyrAlaTyrProPheTyrLysProValAspValGluAlaLeuGluLysHisAspTyr 390
Dy 917 AAGTCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
Qy 391 CysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGluAlaArg 410

```

```

Dy 959 -----AAGTCCCT-----GAGAGGCCAAGTCCCGAGTGAAGGCA--- 994
Qy 411 GLyTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnCysTyr 430
Dy 995 GAGCAAAAGTCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1054
Qy 431 LysTyrAsnProProAspHisGluValValAlaMetAlaArgLysLeuGlnAspValPhe 450
Dy 1055 AAGGCCAAGTCCCGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
Qy 451 GluMetArgPheAlaGluMetProAspGluProGluGluProVal----- 465
Dy 1115 AAGGAGAG---GAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1171
Qy 466 -----ValAlaValSerSerProAlaValProProThrThrLysValAlaProPro 483
Dy 1172 TCCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
Qy 484 SerSerSerAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 503
Dy 1232 GCGCAAGTCCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1282
Qy 504 GLuGluGluArgAlaGlnArgLeuAlaGluLeuGlnGluGlnLysAlaValHisGlu 523
Dy 1283 GAAAAGGAGCAAAAGCCTGTCAGAGAGAGAGTCAAGTCCCGAGAGAGGAGGAGGAGGAGGAGGAG 1342
Qy 524 GlnLeuAlaAlaLeuSerGlnProGlnGlnAsnLysProLysLysGluLysAsp--- 542
Dy 1343 CTGAAGGCGGATGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1402
Qy 543 -----LysLysGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 560
Dy 1403 CCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1462
Qy 561 LysLysSerLysAlaLysGluProProProLysLysThrLysLysAsnSerSerAsn 580
Dy 1463 GAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510
Qy 581 SerAsnValSerLysLysGluProAlaProMetLysSerLysProProProThrTyrGlu 600
Dy 1511 -----ACCAAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1561
Qy 601 SerGluGluGlu-----AspLysCysLysProMetSerTyrGluGluLysArgGln 617
Dy 1562 GAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1621
Qy 618 LeuSerLeuAspIleAsnLysLeu-----ProGluGluLysLeuGluArgValValHis 635
Dy 1622 GAGGCTGAAGTGAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1681
Qy 636 IleIleGlnSerArgGluProSerLysLys-----AsnSerAsnProAspGlu 651
Dy 1682 GTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1741
Qy 652 IleGluIleAspPheGluThrLysProSerThrLeuArgGluLeuGluAlaLeuLysCys 671
Dy 1742 GCCAAGGCCAAG---GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1786
Qy 672 HisLeuLeuPheAlaGluGluLysGluThrPheLysLeuArgLys 686
Dy 1787 -----GCAACGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1819

RESULT 12
US-09-808-880-1/c
; Sequence 1, Application US/09808880
; Publication No. US2003027287A1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE

```



```

Db 16749 GGCACACCTCTCAATCAACCCCGCAGATCCACGACCCCACTCGGAT 16690
Qy 488 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 507
Db 16689 GCTCAGACCCAGAACACCGCCGAAACCCACACTGCGCTGCACCGGATCAATCAACA 16630
Qy 508 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 527
Db 16629 CACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16570
Qy 528 LeuSerGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 542
Db 16569 CCTCAGAACCCGCGCAGATCAACACGCTGCGCAGACACGCGGCAACAC 16510
Qy 543 LysLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 558
Db 16509 CGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16450
Qy 559 GlnGlnLysSerLysSerLysSerLysSerLysSerLysSerLysSerLys 578
Db 16449 CCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16390
Qy 579 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 597
Db 16389 GATCAGGCTCCAGACACTGACCTGACACCGCGCGCTGACGACACACCGCG 16330
Qy 598 ThrTyrGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 617
Db 16329 CAACCAACCAATCATCCGCAAGCCCGCGGACCAACCAACCAACCAACCA 16273
Qy 618 LeuSerLeuSerLeuSerLeuSerLeuSerLeuSerLeuSerLeuSerLeu 637
Db 16272 GCGACTCATCAAGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 16246
Qy 638 GlnSerArgGlnProSerLeuLysSerSerProAspGlnGlnGlnGlnGln 657
Db 16245 GCTTCCACTCCACCGCGGT-----AACGCGAGTCTGTAAGCGCTGATCCACCGCAG 16192
Qy 658 -----ThrLeuLysProSerThr 663
Db 16191 GCCCGCCGACGACGACCAAAAGCCGCAACA 16162

RESULT 13
US-09-938-842A-1265
: Sequence 1265, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Wang, Xun
: APPLICANT: Kreps, Joel
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPT300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227, 866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264, 647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300, 111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 1265
: LENGTH: 1140
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-1265

Alignment Scores: 9.08e-06 Length: 1140
Pred. No.: 226.00 Matches: 86

```

```

Percent Similarity: 35.87% Conservative: 27
Best Local Similarity: 27.30% Mismatches: 117
Query Match: 6.21% Indels: 85
DB: 9 Gaps: 13

US-09-700-590A-22 (1-688) x US-09-938-842A-1265 (1-1140)

Qy 191 ProGlnValSerThrVal-----ProAsnThrGlnAlaSerThr-Pr 205
Db 102 CCGTCCACCAACCCCTTATGTTGATTTGTAATCCGCGGACCACTTCGCGGCAACCGCA 161
Qy 205 oProGlnThrGlnThrPro-----GlnProAsnProProProValGlnAlaThrPr 222
Db 162 TCCCAACCCGCAACCTCCACCAATTTCAACGACACCGCGCAATACCAACCAAC 221
Qy 222 oHisProPheProAlaValThrProAsnLeuLeuValGlnThrProValMetThrVala 242
Db 222 GCGTCCACCAACAGACATCACCACTCCAGTACGACACCAACA-----CGAGACT 275
Qy 242 LProGlnProLeu-----GlnThrProProProValProPr 255
Db 276 TCTCCAAACCCCTTACCGCACCGGTGTCTCTCCGCAACCAACACACACTCTCTCCACC 335
Qy 255 oGlnProGlnProProProAlaProAlaProGlnProValGlnSerHisProProLeu1 275
Db 336 AGCATTTACACCAACCACTCCACCACTATTACACCTTCACATACACTCTCTCCACGAC 395
Qy 275 eAlaAlaThrProGlnProGlnLysThrLysGlnValLysArgLysAlaSerThrTx 295
Db 396 TATTACACCAACCACTCCACTA-----GGGAC 422
Qy 295 rThrProThrThrLysAspProIleHisGlnProProSerLeuProGlnProLysTh 315
Db 423 CACACCAACCGGACTCTCTCCAAACCCCTTACCGCACCGCTGTCTCT-----CCACAAAC 479
Qy 315 rThrLysLeuGlnGlnArgGlnSerSerArgProValLysProProLysAspArva 335
Db 480 CACG-----CAACCTCTCCACCAACCACTCCAC 509
Qy 335 LProAspSerGlnGlnHisProAlaProGlnLysSerSerLysValSerGlnGlnLeuL 355
Db 510 ACCACTATCACT-----CCTCGGTGGAAATATGTTCCAAAGATATACCGAAGCTAA 563
Qy 355 sCyAsSerGlnLysLeuLysGlnMetPheAlaLysLysHisAlaAlaTyrAlaTrpPr 375
Db 564 GATATGAGCGGAAATCTTACCAATTAGTACAGGCTTCTTAACCACTGGAAGAGCAAC 623
Qy 375 oPheTyrLysProValAspValGlnAlaLeuGlnLeuHisAspTyrGlyAspIleLe 395
Db 624 A-----TGTGTTCATCATTCG 641
Qy 395 sHisProMetAspSerSerThrLys-----SerLysLeuGlnAlaArgGlnTyr 412
Db 642 AAATGTGTGATCTTATGATGACAGTACTGTTTGTGAATCGGTAGGACCAACGCTTT 701
Qy 412 rArgAspAlaGlnGlnPheGlnLysAspValArgLeuMetPheSerAsnGlyTyrLysTy 432
Db 702 TTCCCTCTCCCAATTTCCG-----ATCTCTTTAACTTGT----- 741
Qy 432 rAsnProProAspHisGlnValValAlaMetAlaArgLysLeuGlnAspValPheGlnMe 452
Db 742 -----GGTCGTAGAGATTCCACAGGCTTTAAGCTG 770
Qy 452 rArgPheAlaLysMetProAspGlnProGlnGlnProValValAlaLysSerSerProAl 472
Db 771 T-----CCGCGTCCATCCACCAATTCACCAACCAAC 803
Qy 472 aValProProProThrLysValValAlaProProSerSerSer 486
Db 804 ACTTCTCTCAAACTTTAAACCGCTCCGCAACAAACA 846

RESULT 14
US-09-887-576-652

```

```

; Sequence 652, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 652
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-652

```

```

Alignment Scores:
Pred. No.: 9,08e-06 Length: 1140
Score: 226.00 Matches: 86
Percent Similarity: 35.87% Conservative: 27
Best Local Similarity: 27.30% Mismatches: 117
Query Match: 6.21% Indels: 85
DB: 10 Gaps: 13

```

US-09-700-590A-22 (1-688) x US-09-887-576-652 (1-1140)

```

Qy 191 ProGlyValSerThrVal-----ProAsnThrThrGlnAlaSerThr-Pr 205
Db 102 CCTCCACACCAACCTTGTATTGTATTGTATCCGGACCACTCCGCCGACCGCA 161
Qy 205 opProGlnThrGlnThrPro-----GlnProAsnProProProValGlnAlaThrPr 222
Db 162 TCCCCCAACCGCAACTCCACCAATTCAACAGCAGCAGCCAGCCAAATGACCAAC 221
Qy 222 ohisProPheProAlaValThrProAspLeuIleValGlnThrProValMetThrValVa 242
Db 222 GCCTCCACACAGACATCATCACCCTCCAGTAGCGACACACCA-----CCAGCACT 275
Qy 242 lProProGlnProLeu-----GlnThrProProProValProPr 255
Db 276 TCTCTCAAAACCTTACCGCCGACCGTGTCTCTCCGCAAAACACACACACCTCTCC 335
Qy 255 oglnProGlnProProProAlaProAlaProGlnProValGlnSerHisProIleI 275
Db 336 AGCTATTACACCAACCTCCACCAAGTATTACACCTCCACTACCTCTCCACACAGC 395
Qy 275 eAlaAlaThrProGlnProValIleThrIleValIleValIleValIleAlaSerThr 295
Db 396 TATTACACCAACCTCCACTCA-----GCGAC 422
Qy 295 rThrProThrThrIleAspProIleHisGluProProSerLeuProProGluProIle 315
Db 423 CACACACCGGACCTTCTCCAAAACCTTACCGCCGACCGTGTCTCTCT-----CCACAAC 479
Qy 315 rThrIleuGlyGlnArgArgIleSerSerArgProValIleProProIleIleVal 335
Db 480 CAGG-----CCACCTCCCTCCACCAACCAACATTCACAC 509
Qy 335 lProAspSerGlnGlnHisProAlaProGluIleSerSerIleValSerGluGlnLeu 355
Db 510 ACCACATCATCCT-----CCTCTGTTGGATATATGTTCCAGAAATGATACCAACTAAA 563

```

```

Qy 355 sCysCysSerGlyIleLeuIleGluMetPheAlaIleLysHisAlaIleTyrAlaTrp 375
Db 564 GATATGTGCCGGAATTCAGTACATTAGTACGGTCTTCAACCACTGGAAGACAGAAC 623
Qy 375 oPheTyrIleProValAspValGlnAlaLeuGlyLeuHisAspTyrCysAspIleIle 395
Db 624 A-----TGTGTCCCAATTCAG 641
Qy 395 shiSPrometAspMetSerThrIleLys-----SerLysLeuGlnAlaArgIle 412
Db 642 AATGTGTCTGATCTTGAAGCAGTACTGTGTGTAATGCGTAGACACCACTTT 701
Qy 412 rArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnCysTyrIle 432
Db 702 TTCCTCTCTCCCAATTCGGC-----ATCTTCTTAAAGTTGT----- 741
Qy 432 rAsnProProAspHisGluValAlaIleMetAlaIleGlyIleGlnAspValPheGlu 452
Db 742 -----GGTCTGAGATTCCACAAAGCTTTAGCTG 770
Qy 452 rArgPheAlaLysMetProAspGluProGluGluProValAlaIleValSerSerPro 472
Db 771 T-----CCCGGTCCATCACCACATCTCACCAACAC 803
Qy 472 aValProProProThrLysValValAlaProProSerSerSer 486
Db 804 ACTTCTCCCAAACTTTAAACCCGCTCCGCCACCAACACACA 846

```

RESULT 15

```

US-09-764-864-373
; Sequence 373, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 373
; LENGTH: 2653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-373

```

```

Alignment Scores:
Pred. No.: 2.46e-05 Length: 2653
Score: 225.00 Matches: 136
Percent Similarity: 30.78% Conservative: 81
Best Local Similarity: 19.29% Mismatches: 197
Query Match: 6.18% Indels: 291
DB: 10 Gaps: 23

```

US-09-700-590A-22 (1-688) x US-09-764-864-373 (1-2653)

```

Qy 50 ThrSerAsnProAsnLysProLysArgGlnThr-----AsnGlnLeu 63
Db 293 ACCTCTGATGATGCTGATGTCACCAAAAGAAAGAAAGAAACATAAGACGACCAATG 352
Qy 64 GlnTyrLeuLeuArgValIleuLysThrIleTyrIleHisGlnPheAlaTrpPhe 83
Db 353 AAATATTATTCAATGTTGTTCTTGAAGCT-----ValAspAlaValLysLeu 382
Qy 84 GlnGlnPro-----ValAspAlaValLysLeu 92
Db 383 CGAGACCGACGGTTACAGCAGAAAGCTTTGACCTATTATGTTTAAACATCAACAAAAG 442
Qy 93 AsnLeuProAspTyrTyrIleLysIleIleLys-ThrProMetAspMetGlyThrIleLys 112
Db 443 GACTATCTGATATTATATAAATCATCTTGAGCCCAATGACCTTGAAAAATATATTAGCA 502

```

```

Qy 112 sarGleuGluAsnAsnTyTyTTPAsnAlaGlnGluCysIleGlnAspPheAsnThrMe 132
Db 503 TAAATCCGCAATGACAATATGCTGGTGAAGAGGAAATGATGAAGACATGAAGCTCAT 562
Qy 132 tPheThrAsnCysTyTyTleTyAsnLysProGlyAspAspIleValIleuMetAlaGluAl 152
Db 563 GTTCCGAAATCCAGGCACTTAAATGAGAGGCGTCCCAAGTTTAAATGATGACATAT 622
Qy 152 aleuGluLysLeuPheLeuGlnLysIleAsnGluLeuProThrGlnGluThrGluIleMe 172
Db 623 CTGGAGAAAGTTACTCAAGAGAAA----- 647
Qy 172 tIleValGlnAlaLysGlyArgGlyArgGlyArgGlyLysGlnThrGlyThrAlaLysProG 192
Db 648 -----ACGAAAGAGCGGGC----- 662
Qy 192 yValSerThrValProAsnThrThrGlnAlaSerThrProProGlnThrGlnThrProG 212
Db 662 ----- 662
Qy 212 nProAsnProProProValGlnAlaThrProHleProPheProAlaValThrProAspLe 232
Db 662 ----- 662
Qy 232 uIleValGlnThrProValMetThrValValProProGlnProLeuGlnThrProProPr 252
Db 663 -----CC 664
Qy 252 oValProProGlnProGlnProProProAlaProAlaProGlnProValGlnSerHisPr 272
Db 665 ACTGCTCGATGATGATGAC----- 683
Qy 272 oProIleIleAlaAlaThrProGlnProValLysThrLysGlyVal-----Ly 289
Db 684 -----ATGCTTCTCCCAAACTCAGCTGAGTAGAAGAGTGSCATTTCTCTAAAAA 736
Qy 289 sarGlyLeuAlaAspThrThrThrThrProThrThrIleAspProIleGlnGluProProSerLe 309
Db 737 ATCAAAATACATGACTCCCAATGACAGCAAACTAAATGAGGCTATGAA----- 785
Qy 309 uProProGlnProLysThrThrLysLeuGlyGlnArgArgLysSerSerArgProValLys 329
Db 786 -GCTGTAAAGAACTATACTGATAGAGGGGTGCCCTCAGTCCCAATTTCTGAGGCT 844
Qy 329 sProProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSerLys 349
Db 845 TCCTCTAGATCTGAGTGGCT----- 866
Qy 349 sValSerGlnLysLeuLysCysCysSerGlyIleLeuLysGluMetPheAlaLysLysH 369
Db 866 ----- 866
Qy 369 sAlaAlaTyAlaTrpProPheTyTyLysProValAspValGluAlaLeuGlyLeuHisAs 389
Db 867 -----GA 868
Qy 389 pTyTyCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGluAl 409
Db 869 CTACTATCTGACTATTAATAAAAGCCATGACATGAAAAAATTCGAAGTACATATATG 928
Qy 409 aArgGluTyArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnCy 429
Db 929 CAACAAGTACCAAGATATGACTCTATGTTGAGAGACTTTGTCAATGATGTTTAATATGC 988
Qy 429 sTyTyLysTyArgAspProAsp-----HisGluVal 439
Db 989 CTGATAGTACATATGAGCGGAGTCTTTGATCTACAAAGATGCTTTGTTCTACACAAGT 1048
Qy 439 lValAlaMetAlaArgLys----- 445
Db 1049 CCTGCTTGAACAACGACAGACCTTGAGAGAGATGAGAGACTCTCATGTCCCAATGTGAC 1108

```

```

Qy 446 -----LeuGlnAspValPheGluMetArgPheAlaLysMetProAspGluProGluG 463
Db 1109 TTTGCTATTCCAGAGCTCTTCCACATCTTTTGTGTCAATGATGATGATGATGATGATGA 1168
Qy 463 uProValValAlaValSerSer-----ProAlaValProPro-----Pr 476
Db 1169 TGAAGGAAGATGCTACAGCCATCTTTAGCAAGAAATTCCTGCTGTGATGCCAATCTTCC 1228
Qy 476 oThrLysValValAlaProProSerSerSerAspSerSerSerAspSerSerAspSe 496
Db 1229 TAACAAA-----CCACCCCTTACATTTGACATATTAGAAGAAATGTTGAAAATTA 1279
Qy 496 r-----AspSerSerThrAspAspSerGluLys-----GluArgAlaG 509
Db 1280 TCGCTACCGTCCGCTGATTATTTTCAAGAGCATATGTTGAAGTATTTGAAACGACAG 1339
Qy 509 nArgLeu-----AlaGluLeuGlnGlnLys 518
Db 1340 AAGGATGAATCGACAGATTCAGAAATATATGAAGATGACGTAGAACTTCACAGCTTTT 1399
Qy 518 uLysAlaValHisGlnGlnLeuAla-----AlaLeuSe 529
Db 1400 TATTAAATTCGTGATGAATCTGCAAAAATGAGAGATTTCTTCCACCGGACTCAG 1459
Qy 529 rGlnProGlnGlnAsnLysProLysLysGlyLysAspLysLysGlyLys----- 546
Db 1460 CTATACCACAAAACATTTGATATGATGATGAGAAAGAGAGAAAGAAATTTGCCAAA 1519
Qy 547 -----LysLysGlnLysHisLysArgLysGlnGluValGlnGluAsnLysLysSerLys 564
Db 1520 AGAAATGAGAGAAATTAACCTAAACGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTGA 1579
Qy 564 sAlaLysGluProProProLysLysThrLysLysAsnAsnSerSerAsnSerAsnValSe 584
Db 1580 AGATTCTCTGTGCTGTGCTGAGGCTCTCAGGCTTATCATGACATACAGCCAGACTGTAG 1639
Qy 584 rLysLys-----GluProAlaProMetLys 592
Db 1640 CTTTAAAAACGACATGTACATGTTGAGATTAAGTATGTTGAACTCGAAGGCCAA 1699
Qy 592 sSerLysProProProThr----- 598
Db 1700 CTTAAACACACATATCTGCTGTGATTTGAAGAAGCTGGAGAGATTCAGCTGTGAATAATG 1759
Qy 599 -----TyArgLysGlnGluAspLysCysLysProMetSerTy 612
Db 1760 GTTGTATGGCTGTTGTTTAAACGAAATGAATTCACCTGGCTACACGAATAT 1819
Qy 612 rGlnGlnLysArgGlnLeuSerLeuAsp-----IleAsnLysLeuProGlyGlnLys---Le 630
Db 1820 TCTAGAAAAGAAAGATTTTAAAGTGAATTTTAAAGTGAATTTTAAAGTGAATTTTAAAGTGA 1879
Qy 630 uGlyArgValVal 634
Db 1880 AGGCAAGTGTGTG 1892

```

Search completed: March 14, 2003, 22:45:21
 Job time : 208 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 18:46:26 ; Search time 2111 Seconds
(without alignments)
5278.302 Million cell updates/sec

Title: US-09-700-590A-22

Perfect score: 3639
Sequence: 1 MSASGPGTRLRLNLPVMDG.....ALCHLLFAEKEKTFKRLKLM 688

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2.1/USFPO.spool/US09700590/runat.14032003.140630.3244/and.query.fasta.1.839
-DB=EST -OPT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09700590 @cgn 1.1.2874 @runat.14032003.140630.3244 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEORDER -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estlin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_estl: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pin: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_tod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	2914	80.1	1716	11	BC000156	BC000156 Homo sapi
2	2898	79.6	1743	11	BC030158	BC030158 Homo sapi
3	2814.5	77.3	1776	11	BC011541	BC011541 Mus muscu
4	1919.5	52.7	2085	11	BC010699	BC010699 Homo sapi
5	1590.5	43.7	1641	11	BC004832	BC004832 Mus muscu
6	1535.5	42.2	2359	11	BC007715	BC007715 Homo sapi
7	1494	41.1	987	14	BC051563	BC051563 AGENCOURT
8	1484.5	40.8	2440	11	BC001885	BC001885 Homo sapi
9	1436	39.5	893	14	BC021977	BC021977 AGENCOURT
10	1408	38.7	933	14	BC0214403	BC0214403 AGENCOURT
11	1408	38.7	951	14	BC046784	BC046784 AGENCOURT
12	1357.5	37.3	925	14	BC045818	BC045818 AGENCOURT
13	1352	37.2	975	14	BC0653187	BC0653187 AGENCOURT
14	1336.5	36.7	815	14	BC0749323	BC0749323 UI-M-FCO-
15	1319.5	36.3	926	14	BC0650970	BC0650970 AGENCOURT
16	1315	36.1	919	14	BC0652326	BC0652326 AGENCOURT
17	1311.5	36.0	931	14	BC0650713	BC0650713 AGENCOURT
18	1311	36.0	899	14	BC047311	BC047311 AGENCOURT
19	1308	35.9	938	14	BC0654353	BC0654353 AGENCOURT
20	1280	35.2	740	14	BC0683041	BC0683041 UI-CF-EN1
21	1284	34.2	1106	13	BM452148	BM452148 AGENCOURT
22	1225.5	33.7	971	14	BC0652321	BC0652321 AGENCOURT
23	1201	33.0	1452	11	BC028126	BC028126 Homo sapi
24	1200	33.0	1538	11	BC017582	BC017582 Homo sapi
25	1195	32.8	1470	11	BC005281	BC005281 Homo sapi
26	1150.5	31.6	919	12	BF180998	BF180998 601805044
27	1149	31.6	1011	11	BC013666	BC013666 Mus muscu
28	1129	31.0	709	13	BI414784	BI414784 602991186
29	1126	30.9	675	12	BF550759	BF550759 UI-R-CO-J
30	1126	30.9	1055	12	BC824322	BC824322 602727345
31	1114	30.6	641	13	BI907231	BI907231 603065284
32	1099	30.2	658	13	BM389207	BM389207 UI-R-DZO-
33	1072	29.5	682	13	BI252631	BI252631 602952637
34	1066	29.3	621	10	AW539257	AW539257 C0117F08-
35	1061	29.2	631	12	BF712714	BF712714 MI-P-H3-a
36	1053	28.9	627	14	BM679991	BM679991 UI-E-EOL-
37	1052	28.9	666	12	BF718610	BF718610 KEST81 no
38	1039.5	28.6	786	13	BI852791	BI852791 603379265
39	1038	28.5	610	10	AW142274	AW142274 EST292510
40	1036	28.5	619	14	BO939643	BO939643 AGENCOURT
41	1027	28.2	789	13	BC963832	BC963832 602831494
42	1018	28.0	589	13	BM226660	BM226660 K0231D10-
43	1018	28.0	615	12	BC931946	BC931946 602409883
44	1002.5	27.5	629	9	AI552951	AI552951 Y661901.Y
45	992	27.3	902	14	BQ232371	BQ232371 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS BC000156 1716 bp mRNA linear HTC 12-JUL-2001
DEFINITION Homo sapiens, Similar to bromodomain-containing 4, clone
IMAGE:2900081, mRNA.
ACCESSION BC000156
VERSION BC000156.1 GI:14704448
KEYWORDS
SOURCE HTC.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1716)
AUTHORS Strausberg, R.
TITLE Direct Submission

JOURNAL

Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulik, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 3 Row: e Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 3115203
This clone has the following problem: frame shifted.

FEATURES

source

1..1716
location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2900081"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NIH MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT 474 a 549 c 427 g 266 t
ORIGIN

Alignment Scores:

Pred. No.: 1,77e-160 Length: 1716
Score: 2914.00 Matches: 548
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.82% Mismatches: 0
Query Match: 80.08% Indels: 0
Gaps: 0

US-09-700-590a-22 (1-688) x BC000156 (1-1716)

QY 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
Db 69 ATGCTCGGAGAGGCGCCCTGGAGCAATGAGAAATCTGCCAGTATGGGGATGGA 128
QY 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
Db 129 CTAGAAACTTCCCAATGCTCAACACAGAGCCCAAGCCCAAGCCCAAGCCCA 188
QY 41 AlaeSerThrAsnProProProGluThrSerAsnProAsnLysProLysArgGlnThr 60
Db 189 GCCAGCAACCAACCCCGCCCGCCAGAGACCTCCAACTTAACAAAGCCCAAGAGAGACC 248
QY 61 AsnGlnLeuGlnTyrLeuLeuArgValValLeuLysThrLeuTyrLysGlnPheAla 80
Db 249 AACCAACTGCAATACCTCTCAGAGTGGTCTCAAGACACTATGAAACACCAAGTTTCA 308
QY 81 TyrProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
Db 309 TGGCTTTCAGACAGCCGTGGATGCCGTCAAGCTGAACCTCCCTGATTAATAAGATC 368
QY 101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGlnAsnAsnTyrTyrTrp 120
Db 369 ATTTAAAGCGCTTAGATATGGGAACATAAAGAGCGCTTGAAACAACTATTACTGG 428
QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140

Db 429 AATCTCAGAGATGATATCCAGGACTTCAACACTATGTTTCAAAATGTTACTATCAAC 488
QY 141 LysProGlyAspAspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLys 160
Db 489 AAGCTTGAGATGACATAGTCTTAAATGGCAGAGCTGTGGAAAAGCTCTTCTTCAAAAA 548
QY 161 IleAsnGluLeuProThrGluGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
Db 549 ATTAATAGACTTACCAACAGAAAGAAACGAGATCATGTATGCCAGCAAAAGAAAGAGA 608
QY 181 ArgGlyArgLysGluThrGlnTyrThrAlaLysProGlyLysThrValProAsnThrThr 200
Db 609 CGTGGAGAAAGAAACAGAGGACGCAAAACCTGGCTTTCCAGGTATCCAAACACAAC 668
QY 201 GlnAlaeSerThrProProGlnThrGlnThrProGlnProAsnProProProValGlnAla 220
Db 669 CAAGCATGACTCTCTCCGAGACCCCAACCCCTCAGCGAATCTCTCTGTGTCAAGGCC 728
QY 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMetThr 240
Db 729 ACGCTCACCCCTTCCCTGCGGTACCCCGGACCTCATGTCAGACCCCTGTATGACA 788
QY 241 ValValProProGlnProLeuGlnThrProProProValProProGlnProGlnProPro 260
Db 789 GTGTGTCTCTCCAGCCACTGACAGACGCCCCGCAAGTGCCTCCAGCAACACCCCA 848
QY 261 ProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaAlaThrProGln 280
Db 849 CCCCTCAGTCTCCCGACGCCGTACAGACCAACCCCATCATCGGGCCACCCCAAG 908
QY 281 ProValLysThrLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle 300
Db 909 CCTGTGAAGACAAAGAGGAGGTGAAGAGAAAGACAACACACACCCCAACACCAT 968
QY 301 AspProIleHisGluProProSerLeuProProGluProLysThrThrLysLeuGln 320
Db 969 GACCCCATTCACGAGCACCTGCTGCTCCCGGAGCCCAAGACCAACAGCTGGGACAG 1028
QY 321 ArgArgLysSerSerArgProValLysProProLysLysAspValProAspSerGlnGln 340
Db 1029 CGGCGGAGAGACACCGGCTGTGAACCTCCAAAGAAAGACGTGCCGACTCCACAG 1088
QY 341 HisProAlaProGluLysSerSerLysValSerGlnLeuLysCysCysSerGlyIle 360
Db 1089 CACCAGACACAGAGAGAGAGAGCAAGGTCTCGGAGCAGCTCAAGTCTGACGCGCATC 1148
QY 361 LeuLysGluMetPheAlaLysLysHisAlaAlaTyrAlaTrpProPheTyrLysProVal 380
Db 1149 CTCAAGAGATGTTTGCCTCAAGAAACAGCGCTTACGCTTCTTCAACAGCTGTG 1208
QY 381 AspValGlnAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisProMetAspMet 400
Db 1209 GACGTGAGGCACTGGGCTTACAGACTACTGTGACTCATCAAGCAACCCATGAGCATG 1268
QY 401 SerThrIleLysSerLysLeuGlnAlaArgGluTyrArgAspAlaGlnGluPheGlyAla 420
Db 1269 AGCAATCAAGTCTTAATCTGAGAGCCCGTAGTACCTGATGCTCAGAGATTGTGCT 1328
QY 421 AspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHisGlnValVal 440
Db 1329 GACGTCCGATTGATGTTCTCCAACTGCTATAGTACAACCTCTGACATGAGAGTGTG 1388
QY 441 AlaMetAlaArgLysLeuGlnAspValPheGluMetArgPheAlaLysMetProAspGlu 460
Db 1389 GCCATGGCCGCAAGCTTCCAGAGATGTGTGAATAGCGCTTTCAGAAAGATCCGGAAGAG 1448
QY 461 ProGluGluProValValAlaValSerSerProAlaValProProProThrLysValVal 480
Db 1449 CCTAGAGAGCAGTGTGTGCGGTCTCTCCCGGCAATGCCCCCTCCACCAAGTTGTG 1508
QY 481 AlaProProSerSerSerAspSerSerSerSerSerSerSerSerSerSerSerThr 500
Db 1509 GCCCCGCTCATCTCAGAGAGAGAGAGAGATGCTCTCGAGCAAGTGAAGTTGACT 1568

Db 1059 CGGGCGAGACGACCGGCTGTGAAACCTCCAAAGAGACGCGGCTTCAGCAG 1118
 Qy 341 HisProalPaProgluLysSerSerLysValSerGluLysCysCysSerGlyLe 360
 Db 1119 CACCCAGACACAG 1118
 Qy 361 LeuLysGluMetPheAlaLysLysHisAlaAlaAlaAlaAlaAlaAlaAlaAla 380
 Db 1179 CTCAAGAGAGATGTTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238
 Qy 381 AspValGluAlaLeuGlyLeuHisAspValPheGluMetArgPheAlaLysMet 400
 Db 1239 GACGTGGAGGACCTGGGCTCAGACACTGATGATCATCATCAACACCCCACTGAG 1298
 Qy 401 SerThrIleLysSerLysLeuGluAlaArgGluValAspAlaGluLysPheGlyAla 420
 Db 1299 AGCACAATCAAGCTCAAACTGAGAGCCCGCTGAGTACCTGATGCTCAGAGAGTTG 1358
 Qy 421 AspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHisGluValAla 440
 Db 1359 GACGTCCGATTGATGTCTCCAACTGATATAGTCAACCTCCTGAGCAGATGAGTGTG 1418
 Qy 441 AlaMetAlaArgLysLeuGluAspValPheGluMetArgPheAlaLysMetProAspGlu 460
 Db 1419 GCCATGGCCCAAGCTCAGAGATGTTCGAAATGCGCTTGGCCAAATGCGGAGACAG 1478
 Qy 461 ProGluGluProValAlaAlaValSerSerProAlaValProProThrLysValAla 480
 Db 1479 CCGAG 1538
 Qy 481 AlaProProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerThr 500
 Db 1539 GCCCGGCGCTCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
 Qy 501 AspAspSerGluGluGluValArgAlaGluArgLeuAlaGluLeuGluGluLysAla 520
 Db 1599 GATGACTCTGAG 1658
 Qy 521 ValHisGluGluLysAlaAlaLeuSerGluProGluGluLysAlaProLysLysGlu 540
 Db 1659 GTGACGAG 1718
 Qy 541 LysAspLysLysGluLysLysLys 548
 Db 1719 AAAAAAAAAAAAAAAAAAAAAA 1742
 RESULT 3
 BC011541
 LOCUS BC011541 1776 bp mRNA linear HTC 07-AUG-2002
 DEFINITION Mus musculus, similar to bromodomain-containing 4, clone
 IMAGE:2648506, mRNA.
 ACCESSION BC011541 GI:15079397
 VERSION BC011541.1
 KEYWORDS HTC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1776)
 TITLE Direct Submission
 JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcaps-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 5 Row: J Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10048443
 This clone has the following problem: incomplete processing.
 Location/Qualifiers
 1..1776
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="IMAGE:2648506"
 /issue_type="Mammary tumor. Metallothionin-TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_lib="NCI CGAP Mam1"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 BASE COUNT 545 a 517 c 401 g 313 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,13e-154 Length: 1776
 Score: 2814.50 Matches: 531
 Percent Similarity: 96.59% Conservative: 7
 Best Local Similarity: 95.33% Mismatches: 18
 Query Match: 77.34% Indels: 1
 Gaps: 1
 DB: 11
 US-09-700-590a-22 (1-688) x BC011541 (1-1776)
 Qy 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 Db 102 ATGTCTACGAG 161
 Qy 21 LeuGluThrSerGluMetSerThrThrGluAlaGluAlaGluProGluProAlaAsnAla 40
 Db 162 CTAGAAACCTCCCAATATGTCTACAAACGAGAGAGAGAGAGAGAGAGAGAGAGAG 221
 Qy 41 AlaSerThrAsnProProProGluThrSerAsnProAsnLysProLysArgGluThr 60
 Db 222 GCCAGACCAATCTTCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 281
 Qy 61 AsnGluLeuGluThrLeuLeuArgValValLeuLysThrLeuTyrPheLysPheAla 80
 Db 282 AACCACTGCAATATCTGCTCAGAGTGTGTCTCAAGACATATGAGAAACACAGTTTGGC 341
 Qy 81 TyrProPheGluGluProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysLe 100
 Db 342 TGGCCTTCCAG 401
 Qy 101 IleLysThrProMetAspMetGlyThrLileLysLysArgLeuGluAsnLysTyrTyr 120
 Db 402 ATTAACAACCCATGATGATGAGAAACATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 461
 Qy 121 AsnAlaGluGluCysIleGluAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140
 Db 462 AATGCTCAGGAATATATCCAGAGACTTCAACACTATGTTTCAAAATGTTCATATTAAC 521
 Qy 141 LysProGluAspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGluLys 160
 Db 522 AAGCTGAGATGACATCGCTTATATGAGCAAGAGCTCGGAGAGAGCTCTTCTTCAAAAA 581
 Qy 161 IleLeuGluLeuProThrGluGluThrGluLysLeuMetIleValGluAlaLysGlyArgGly 180
 Db 582 ATCAATGAATGCTTACAGAGAAATGAGATCATATGATGTCAGGCAAAAGAGAGAGAG 641

[illegible][illegible]

QY 64 GlnTyrLeuLeuValValLeuLysThrLeuTyrLysHisGlnPheAlaTyrProPhe 83
 |||||
 Db 276 CAGTACATGACGAGATGTGTGTGTAAGACGCTGGAACACCAAGCTTCGCTGGCCCTTC 335
 QY 84 GlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysLysLysThr 103
 |||||
 Db 336 TACACGCCCGGAGCGCAATTAATTAACCTGCGGATTCATTAATAATATTAATAAAC 395
 QY 104 ProMetAspMetGlyThrLysLysLysArgLeuGlnAsnTyrTyrTyrAspAlaGln 123
 |||||
 Db 396 CCAATGATATAGGAGACTATTAAGAGAGACTAGAAATATATATATATATGAGTGAAGC 455
 QY 124 GluCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrTyrTyrAsnLysProGly 143
 |||||
 Db 456 GAATGTATGACGAGACTTCAACACCATGTTTCAAAATGTTTCAATTATTAACAGCCCA 515
 QY 144 AspAspIleValLysLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLysLys 163
 |||||
 Db 516 GATACATAGTGTAAATGCGCCAGCTTAAAGAAATTTTTCACAAATAATGAGCCGAG 575
 QY 164 LeuProThrGlnGluThrGluLysLysLysLysLysLysLysLysLysLysLysLys 181
 |||||
 Db 576 ATGCCCCAAGAGAGAGTGAATTAATTAACCCCTGCTCCAAAGCGCAAGTCCGAAAGCG 635
 QY 182 -----GlyArgLysGluThrGlyTyrLysLysProGlyValSerThrValProAsnThr 199
 |||||
 Db 636 GCTCGGAGCGCCAGAGCGCAGGTACAGCA-----GTGGCGCGCGTGTCTCTCTC 689
 QY 200 ThrGlnAlaSerThrProProGlnThrGlnThrPro-----GlnProAsnProProVal 218
 |||||
 Db 690 TCCCGAGG-----ACCCCTTTTCAGCGCTGCGCCCGCCACCTCTC 728
 QY 219 GlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProVal 238
 |||||
 Db 729 TCCCAAGAGCGCC-----GTCTGCTGCGCCACCCCTCTTA 761
 QY 239 MetThrValValProProGlnProLeuGlnThrProProProValProProGlnProGln 258
 |||||
 Db 762 CCAACCATCAGTCGCAAC-----GTCAAGTGTGTCCAGTCCCGCCAGCTGCGCC 812
 QY 259 ProProProAlaProAlaProGlnProValGlnSerHisProProIleIleAlaIleThr 278
 |||||
 Db 813 CCACCTCTCT-----CTGCAACACCATCTGCTGCTGCTCTCTCTG 857
 QY 279 ProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThrThrProThr 298
 |||||
 Db 858 CCG-----CTGTGTCTCAAGAAAGAGCGGTGAAGCGAAAGCAACACCATCTCCACG 914
 QY 299 -----ThrIleAspProIleHisGluProProSerLeuProProGlnProLys 314
 |||||
 Db 915 ACGTGGCGCATCACTGCCAGCGGAGTGTGCTGCGCCCGCTG---TCAGACCCCAAG 971
 QY 315 ThrThrLysLeuGlyGlnArgArgLysSer---SerArgProValLysProProLysLys 333
 |||||
 Db 972 CAGCGCAAAAGTGTGCGCGCGGAGAGTGTGCGCGCCCATCAACCTCTCCCAAGAG 1031
 QY 334 AspValProAspSerGlnGlnHisProAlaProGluLysSerSerLysValSerGln 353
 |||||
 Db 1032 GACCTGGAGAGAGCGGAGGTGCGCCAGCAGCAGCAGCAGCAGCAGCAGTGTGCGAGC 1091
 QY 354 LeuLysCysCysSerGlyLysLeuLysGluMetPheAlaLysLysHisAlaLysLys 373
 |||||
 Db 1092 CTACGCTACTGAGACAGATCTCTGAGGAGAGTATCAAAAGACACGCGGCTTAAGCC 1151
 QY 374 TyrProPheTyrLysProValAspValGluAlaLeuGlyLysHisAspTyrCysAspIle 393
 |||||
 Db 1152 TGGCCCTTCTCAAGCCAGTGTGATGCGCGCGCTGTGAGCTGCAGACTACCAACACATTC 1211
 QY 394 IleLysHisProMetAspMetSerThrLysSerLysLysGluAlaArgLysLysLys 413
 |||||
 Db 1212 ATCAAGACCCGATGAGCTCAAGACCGTGAAGAGAGATGAGTGGCGAGAGTACCA 1271

QY 414 AspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnCysTyrLysTyrAsn 433
 |||||
 Db 1272 GAGCGACAGGCGCTTGTGTGTATGTCCGACTGATGTTCTCGAATTCCTCAATCAAT 1331
 QY 434 ProProAspHisGlnValValAlaMetAlaArgLysLeuGlnAspValPheGluMetArg 453
 |||||
 Db 1332 CCCCACACACAGAGGTGTGTGCGCATGTGCGCGAGCTCCAGAGCGTGTGAGATGAG 1391
 QY 454 PheAlaLysMetProAspGluProGluGluProValValAlaValSerSerProAlaVal 473
 |||||
 Db 1392 TTGGCCAAAGATGCGCAGATGAGCC-----GTGAGCGCACCGCGCTG 1433
 QY 474 ProProThrThrLysValAlaProProSerSerSerAspSerSerSerSerSer 493
 |||||
 Db 1434 CTGCCCCCGCGCGCCCATGTGTGAGCAGAGCGGCTGTAGACGACCTTACAGTGAAG 1493
 QY 494 SerAspSerAspSerSerThrAspAspSerGluGluGluArgAlaGlnArgLeuAlaGlu 513
 |||||
 Db 1494 AGCTCTTGAGACTAGCAGCAGCTCGGACTCGGAGAGAGAGCGGCGCACAGGCTGCGAG 1553
 QY 514 LeuGlnGluGlnLeuLysAlaValHisGluGlnLeuAlaLeuSerGlnProGlnGln 533
 |||||
 Db 1554 CTGCAAGAGAGCTGAAGCGCGGTGCAAGACAGCTGCGCCCTGTCTCAGGCCCCAGTA 1613
 QY 534 AsnLysProLysLys-----LysGluLysAspLysLysGluLysLys 547
 |||||
 Db 1614 AACCAACCAAGAAAGAAAGAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1673
 QY 548 LysGluLysHisLysArgLysGluGluValGluGluLysLysLysLysLysLysLys 564
 |||||
 Db 1674 AAGGAG 1733
 QY 565 ---AlaLysGluProProProLysLysThrLysLysAsnSerSerSerSerSerSer 583
 |||||
 Db 1734 CCTGCCAAGCAGCTCAGCAGAGAGAGAGAGCTCTGCCAAGAGCGCAACACACACAGC 1793
 QY 584 SerLysLysGluProAlaProMetLysSerLysProProProThrTyrGlnSerGluGlu 603
 |||||
 Db 1794 GCCCGACAGCAGCTGAAGAAAGAGCGGAGAGAGAGATCTGCTCTCACTCAGAGAGAA 1853
 QY 604 GluAspLysCysLysProMetSerTyrGluGluLysArgLysLeuSerLeuAspIleAsn 623
 |||||
 Db 1854 GAGAGAGAGAGCGCTCCCATATGAGTACATGAAGAGCGCAGCTTACCTGGAATCAAC 1913
 QY 624 LysLeuProGlyGluLysLysGlyArgValHisLysLysLysLysLysLysLysLys 643
 |||||
 Db 1914 CGGCTGCCGCGGAGAGAGCTGCGCGGAGTGTGCAATCCAAATCTCGGAGCCCTCG 1973
 QY 644 LeuLysAsnSerAsnProAspGluLysLysLysLysLysLysLysLysLysLysLys 663
 |||||
 Db 1974 CTCAAGGCTTCAACCCCGAGAGATGAATTAATTTGAGACTGTGAACCCACACT 2033
 QY 664 LeuArgGluLeuGlyAlaLeuLysCysHisLeuLeuPheAlaGluLysGlu 680
 |||||
 Db 2034 TTGGCGGAACT-GGAGAGATATGTCAAGTCTTTCACAAAAAATAAAAAA 2083
 RESULT 5
 BC004832 1641 bp mRNA linear HTC 07-AUG-2002
 LOCUS BC004832
 DEFINITION Mus musculus, similar to RIKEN cDNA 2410084F24 gene, clone
 IMAGE:2650994, mRNA.
 ACCESSION BC004832
 VERSION BC004832.1 GI:14709507
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1641)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer

Db 1398 -----GCACCTGCGCTGCGCTCCCGCCAGCCCGCCCATCTGAGCAG 1439

Qy 484 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 503

Db 1440 GGGGCTGAGAGACCGGACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499

Qy 504 GUG 523

Db 1500 GAG 1559

Qy 524 GUG 543

Db 1560 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616

Qy 544 LysGluLysLysLysGluLys 550

Db 1617 AAGGAAAAAGAAAAAAA 1637

RESULT 6
BC007715 2359 bp mRNA linear HTC 12-JUL-2001
LOCUS Homo sapiens, bromodomain-containing 2, clone IMAGE:4302993, mRNA.
DEFINITION BC007715
ACCESSION BC007715 GI:14712779
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgrl.nih.gov

Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Boulford, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Masileilo, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Starick, P., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAL Plate: 17 Row: 0 Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12408641
This clone has the following problem: incomplete processing.

FEATURES

SOURCE

1. 2359
/organism="Homo sapiens"
/db_xref="locusID:6046"
/db_xref="taxon:9606"
/clone="IMAGE:4302993"
/tissue_type="uterus", leiomyosarcoma"
/clone_id="NTH_MGC_46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 603 a 683 c 593 g 480 t
ORIGIN

Alignment Scores:
Pred. No.: 4.24e-80 Length: 2359
Score: 1535.50 Matches: 319
Percent Similarity: 66.61% Conservative: 62
Best Local Similarity: 55.77% Mismatches: 120
Query Match: 42.20% Indels: 71
DB: 11 Gaps: 13

US-09-700-590a-22 (1-688) x BC007715 (1-2359)

Qy 5 SerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGluThrSer 24
Db 770 GCAGCACCAGGAGAAAGATTCGAAACCTCTCTCTGATATAGAGGCTTGAGAGGCC 829

Qy 25 GlnMetSerThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaAlaSerThrAsn 44
Db 830 ACATGGCTTCGGTGGCTGCT--TTGCACTTAACCTGCTC-----AAC 871

Qy 45 ProProProProGluThrSerAsnProAsnLysProLysArgGlnThrAsnGlnLeuGln 64
Db 872 CCACACACCCCGGAGGTGTCCAAATCCCAAAAGCAGAGAGATTACCAACAGCTGCAA 931

Qy 65 TyrLeuLeuArgValValLeuLysThrLeuTyrLysHisGlnPheAlaTyrProPheGln 84
Db 932 TACTTACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 991

Qy 85 GlnProValAspAlaValLysLysLeuAsnLeuProAspTyrTyrLysIleIleLysThrPro 104
Db 992 CAGCTGTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051

Qy 105 MetAspMetGlyThrIleLysLysArgLeuGlnAsnLysTyrTyrTyrPheAlaGlnGlu 124
Db 1052 ATGACACATGTTGATATTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111

Qy 125 CysIleGlnAspPheAsnThrMetPheThrAsnGlyTyrIleTyrAsnLysProGlyAsp 144
Db 1112 TGTATGACAGATTTATTAATCATTTTCAACACTGTTTACATTTCACCAAGCCACATGAT 1171

Qy 145 AspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLysIleAsnGlnLeu 164
Db 1172 GATTTTGCTCAATGACGACAAAGCTGGAAGATATTCACAGAGAGGTTGATCATATG 1231

Qy 165 ProThrGluGluThrGluIleMetIleValGlnAlaLysGly-----ArgGly--- 180
Db 1232 CCACAGAGAGAAAGAGAGCTGTAGTACCATCTTAAGAACAGCCCAAGAGAGGAGCC 1291

Qy 181 -----ArgGlyArgLysGluThrGlyThrAlaLysProGlyValSerThr 195
Db 1292 AAGTTGGACGCTCCAGGAGAGTTCACAGTCCATCAGTGCCTGCTGCTCTTCT 1351

Qy 196 ValProAsnThrThrGlnAlaSerThrProProGlnThrGlnThrProGlnProAsnPro 215
Db 1352 GTTCACACACAGAGCCGTATATCTCTCCACCTGAGATA-----CCT 1393

Qy 216 ProProValGlnAlaThrProHisProPheProAlaValThrProAspLeuIleValGln 235
Db 1394 ACCATGCTCCCAACATTCCTCCACCA-----TCAGTATTTCC 1432

Qy 236 ThrProValMetThrValAlaProProGlnProLeuGlnThrProProValProPro 255
Db 1433 TCTCCACTTCTCAAGTCTTGCAC-----TCTGCTGACCCCGCTCTTGTCT 1480

Qy 256 GlnProGlnProProProAlaProAlaProGlnProValGlnSerHisProProIleIle 275
Db 1481 GTTACTGAGCTCTCTCAAGC----- 1501

Qy 276 AlaAlaThrProGlnProValLysThrLysLysGlyValLysArgLysAlaAspThrThr 295
Db 1502 -----CAGCCCTTGCCAGAGAAAAAGCGCTAAAGCGGAAAGCAGATATACAC 1549

Qy 296 ThrProThrThrIle-----AspProIleHisGlnProProSerLeu 309

Db 1550 ACCCTACACCTACAGCATCTTGGCTCTGTTCTCCACTACCTCTCTGGAGCTT 1609
 Qy 310 ProProGluProLyThrThrLySleu--GlyGlnArgArgLysSerArgProVal 328
 Db 1610 -----GAGCTTAAGCAGACGAGCTTCCCCCTATGCTAGAGAGATGTCGCCCATC 1663
 Qy -329 LysProProLySlyAspValProAspSerGlnGlnHisProAlaProGluLySerSer 348
 Db 1664 AACCCCCCAGCAAAACCTTGGCTGACTTCACACACACACAGAGCTTAAGAAAGA 1723
 Qy 349 LysValSerGluGlnLeuLySlyCysSerGlyLlleuLySgluMetPheAlaLySly 368
 Db 1724 AGCTTTCAGAACAGTTAAACATTGCATGCGATTTTGAAGAGATTACTCTTAAGAG 1783
 Qy 369 HisAlaAlaTyrAlaTyrProPheTyrLySProValAspValGluAlaLeuGlyLeuHis 388
 Db 1784 CAGTGGCTCATGCTTGGCTTCTTATTAACCAAGGATGATCTTGGACACTTGGCTGCAT 1843
 Qy 389 AspTyrCysAspIleIleLySHisPheMetAspMetSerThrIleLySerLySleuGlu 408
 Db 1844 GACTACCATGACATCATTTAAGACACCCCATGAGCTCAGCACTGTCAAGCGAGATGGAG 1903
 Qy 409 AlaArgGluTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsp 428
 Db 1904 AACCGTATTAACCGGAGATGACAGAGTTTCTGCTGATGTACGGCTTATGTTCTCAAC 1963
 Qy 429 CysTyrLySlyTyrAspProAspHisGluValValAlaMetAlaArgLySleuGlnAsp 448
 Db 1964 TGCTATTAAGTACATCCCCAGATCAAGATGTTGTGGCAATGGCAGAAAGTACAGAT 2023
 Qy 449 ValPheGluMetCysPheAlaLySlyMetProAspGluProGluGluProValValAlaVal 468
 Db 2024 GATTTTGAAGTTCGGTATATGCAAGATGCGAGATGAACCATTAACACGAGGCTTTACCA 2083
 Qy 469 SerSerProAlaValProProPro---ThrLySValValAlaProProSerSerSerAsp 487
 Db 2084 GTCCTTACTGCAATGCGCCCTGGCTGGCCAAATCGTCTTCAAGATCTTCACTAGAGAA 2143
 Qy 488 SerSerSerAspSerSerSer-----Asp 495
 Db 2144 AGTACAGATGAGAGCTCTCTGAGAAAGAGAGAGAGATAGAGAGACGAGAGAGAA 2203
 Qy 496 SerAspSerSerThrAspSerSerGlnGluArgAlaGlnArgLeuAlaGluLeuGln 515
 Db 2204 GAAGAGAGTGAAGCTCAGACTCAGAGAAAGAGGCTCATCGCTTACAGAACTACAG 2263
 Qy 516 GlnGlnLeuLySlyAlaValHisGluGlnLeuAlaLeuSerGlnProGlnGlnAsnLyS 535
 Db 2264 GAACAGCTTCGGGCAAGTACATGAACAACCTGCTGCTGCCAGGGGTCCAAATATCCAA 2323
 Qy 536 ProLySlySlyGluLySAspLySlySgluLySlyS 547
 Db 2324 CCCAAGAGAAAGAGAGAAAAA 2359
 RESULT 7
 LOCUS BQ651563 987 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT B302486 NIH_MGC_100 Homo sapiens CDNA clone IMAGE:6271145
 5' mRNA sequence.
 ACCESSION BQ651563
 VERSION BQ651563.1 GI:21775735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 987)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@dbp-r@mail.nih.gov
 Tissue Procurement: CGAP (Stanford)

cDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LICH2447 row: g column: 18
 High quality sequence start: 2
 High quality sequence stop: 653.
 Location/Qualifiers
 1..987
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6271145"
 /clone_1b="NIH MGC 100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="RD10B (phage-resistant)"
 /note="organ: liver; Vector: pOTB7, site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dt priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-CDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH MGC
 Library."

BASE COUNT 257 a 315 c 267 g 136 t 12 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,6e-78 Length: 987
 Score: 1494.00 Matches: 288
 Percent Similarity: 96.37% Conservative: 4
 Best Local Similarity: 95.05% Mismatches: 11
 Query Match: 41.06% Indels: 1
 DB: 14 Gaps: 0

US-09-700-590A-22 (1-688) x BQ651563 (1-987)
 Qy 250 ProProProValProProGlnProGlnProProProProAlaProAlaProGlnProValGln 269
 Db 9 CCTGGGTGGCTGCCCCCGCCAGCACACACCCCGGCTCCAGCTCCCGAGCCGACAG 68
 Qy 270 SerHisProProIleIleAlaAlaThrProGlnProValLyThrLySlySlyValLyS 289
 Db 69 AGCCACCCACCATCATCGCGGCCACCCACAGCTGTGAACAACAAGAGAGAGTGAAG 128
 Qy 290 ArgLySAlaAspThrThrThrProThrThrIleAspProIleHisGluProProSerLeu 309
 Db 129 AGGAAAGCAGACACACACACCCACACCATGACCCCATTCACGAGCCACCTCGCTG 188
 Qy 310 ProProGluProLyThrThrLySleuGlyGlnArgArgLysSerSerArgProValLyS 329
 Db 189 CCCCCGAGCCCAAGACCAACCAAGCTGGGCGCAGCGGAGAGAGAGCGCGCTTGAAA 248
 Qy 330 ProProLySlyAspValProAspSerGlnGlnHisProAlaProGluLySerSerLyS 349
 Db 249 CCTCCAAAGAAAGAGTGGCGGACCTCTACAGCAGACCCAGACCAAGAGAGAGCAGAG 308
 Qy 350 ValSerGluGlnLeuLySlyCysSerGlyLlleuLySgluMetPheAlaLySlyHis 369
 Db 309 GTCGCGAGCACTCAAGTGTGCGAGCGGCATCCCAAGAGATGTTGCCAAGAGCAC 368
 Qy 370 AlaAlaTyrAlaTyrProPheTyrLySProValAspValGluAlaLeuGlyLeuHisAsp 389
 Db 369 GCGGCTTACGCTGGCCCTTCTCAAGCTTGTGACGTGAGAGCACTGGGCTTACAGCAG 428
 Qy 390 TyrCysAspIleIleLySHisPheMetAspMetSerThrIleLySerLySleuGluAla 409
 Db 429 TACTGTGACATATCAAGACCCCATGACATGACACACATCAATCTTAATCGAGGCC 488
 Qy 410 ArgGluTyrArgAspAlaGlnGluPheGlyAlaAspValArgLeuMetPheSerAsnLyS 429

Db	489	CGTAGACACCGTAGTATGCTCAGGAGTTTGCTGACCGTCGATTGATGTCTCCAACTGC	548
QY	430	TyrlsYfsYrAnP-ProP-ROAphSiGLuValAlaMetAlaArgLysLeuGlnAspVal	449
Db	549	TATTAAGTACAACCTCTCCCTGCACATGAGGTGTGGCCACAGGCCCAAGCTCCAGATGTG	608
QY	450	PheGluMetArgPheAlaYMetProAspGluProGluGluProValAlaValSer	469
Db	609	TTCCAAATGCACCTTTGCCAAGATGCCGACAGACTTGAGAGCCAGTGTGGCCGTGCC	668
QY	470	SerProAlaValProProProThrLysValValAlaProProSerSerSerSerSer	489
Db	669	TCCTCCGGAGTGGCCCTCCACCAAGGTGTGGCCCGCCCTCATCCAGCAGACGACG	728
QY	490	SerAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	509
Db	729	AGCCATAGCTCTCCGACAGTACGACGTTCGACTGATGCTCTGAGAGACGACGCCAG	788
QY	510	ArgLeuAlaGluLeuGlnGluGluLeuLysValAlaHisGluGlnLeuAlaLeuSer	529
Db	789	CGGCTGGCTGAGCTCAGGAGACAGCTCAAGCCCTGCACGACGACCTTGCAAGCTCT	847
QY	530	GlnProGlnGlnAsnLysProLysLysLysGluLysAspLysGluLysLysGlu	549
Db	848	CAGCCCGACGACGACCAACCAAAAAAAAAAAAAAAAAAGGANNNNNNANANAAGAA	907
QY	550	LysHisLys 552	
Db	908	AGAAAAA 916	
RESULT 8			
LOCUS	BC001885	2440 bp	mRNA
DEFINITION	Homo sapiens, similar to bromodomain-containing 2, clone		HTC 12-JUL-2001
ACCESSION	IMAGE:3544244, mRNA.		
VERSION	BC001885.1	GI:14706221	
KEYWORDS	HTC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2440)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JAN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D. cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca		

Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywnicki, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven
Nees, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Diane Smilans, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 8 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4757809
This clone has the following problem: frame shifted.

FEATURES		Location/Qualifiers
source	1..2440	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3544244" /tissue_type="Lymph, Burkitt lymphoma" /clone_id="NH MGCC 8" /lab_host="DH10B-R-" /note="Vector: pOTB7"
BASE COUNT 619 a 710 c 609 g 502 t		
ORIGIN		
Alignment Scores:		
Pred. No.:	4,06e-77	Length: 2440
Score:	1484.50	Matches: 318
Percent Similarity:	63.02%	Conservative: 62
Best Local Similarity:	52.74%	Mismatches: 121
Query Match:	40.79%	Indels: 103
DB:	11	Gaps: 14
US-09-700-590A-22 (1-688) x BC001885 (1-2440)		
QY	5 SerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGlyLeuGlnThrSer	24
DB	759 GCACGACCAAGGAAAGATTGCAAAACCTCTCTTGATGAGGGCTTGAAGAGCC	818
QY	25 GlnMetSerThrGlnAlaGlnAlaGlnProGlnProAlaAsnAlaLaserThrAsn	44
DB	819 ACAATGGCTTGGTGGCTGGT---TTGCAACTTACCCCTGCC-----AAC	860
QY	45 ProProProGlnThrSerAsnProAsnLeuProLysArgGlnThrAsnGlnLeuGln	64
DB	861 CCACCACCCCGAGGCTGTCATTCACAAAAGCCAGAGACGATTAACACACAGCTGAA	920
QY	65 TyrLeuLeuArgValValLeuLysThrLeuTrpLysGlnPheAlaTrpProPheGln	84
DB	921 TACCTTACAAAGGATGATGAAGAGCTCTGTGAAACAATCAGTTGCGATTCGCG	980
QY	85 GlnProValAspAlaValLysLeuAsnLeuProAsp-----	96
DB	981 CAGCTGTGATGCTGTCAAACTGGGTCTACCGAT-TCCACCTTGGGTTGGAGAGAC	1039
QY	96 -----	96
DB	1040 CACGGTGCCAAATTTCTTACGCTTCTTCTTCCTCATGACGCCCATGATAGCCACC	1099
QY	97 -----TyrTyrLysIleIleLysThrProMetAspMetGlyThrIleLysLysArg	113
DB	1100 CCAGAGATTATACAAATATTATTAACACGCTATGACATGGTACTTATTAAGAGAGA	1159
QY	114 LeuGlnAsnAsnTyrTyrTrpAsnAlaGlnGluCysIleGlnAspPheAsnThrPhe	133
DB	1160 CTTGAAAAACAATTATTAATTGGGCTGTCTCAGAGTGTATGCCAAGATTTTAATACATGTC	1219
QY	134 ThrAsnCysTyrIleTyrAsnLysProGlyAspAspIleValLeuMetAlaGlnAlaLeu	153
DB	1220 ACCAAGCTTATCATTTTCAACAAAGCCCACTGATGATATTGTCTTAATGGCACAAACGCTG	1279
QY	154 GlnLysLeuPheLeuGlnLysIleAsnGlnLeuProThrGlnGluThrGluIleMetIle	173
DB	1280 GAAAGATATTCTTCACGAAGAGTTGGATCAATGSCCAAGAAAGAACAGAGCTGGTACTG	1339
QY	174 ValGlnAlaLysGly-----ArgGlyLys	184
DB	1340 ACCATTCCTTAAGAACACGACCAAGAGGGGGCCAAAGTGGCAGCGCTCCAGGGCAGTGT	1399
QY	185 GlnThrGlyThrAlaLysProGlyValSerThrValProAsnThrThGlnAlaSerThr	204
DB	1400 ACCAGTGCATCAGGCTGCGCTCTCTTGTGTCAACACAGCCCTGTATATCTCT	1458
QY	205 ProProGlnThrGlnThrProGlnProAsnProProValGlnAlaThrProIleAsp	224
DB	1460 CCACCTGAGATA-----CTTACACAGCTGTCTCAACATTTCCCAACCA	1501


```

QY 225 PheProAlaValThrProAspLeuIleValGlnThrProValMetThrValValProPro 244
DB 1502 -----TCAGTCATTTCTCTCCACTCTCAAGTCTTGACAC--- 1537
QY 245 GlnProLeuGlnThrProProProValProProGlnProGlnProProProAlaProAla 264
DB 1538 -----TCTGTGAGACCCGCTCTGTACTGACTGAGCTCTTCAGCC--- 1582
QY 265 ProGlnProValGlnSerHisProProIleIleAlaAlaThrProGlnProValIleThr 284
DB 1583 -----CAGCCCTTGCCAG 1597
QY 285 LysLysGlyValLysArgLysAlaAspThrThrThrProThrThrIle----- 300
DB 1598 AAAAAAGGCGTTAAAGCGAAGACAGATACACACCCCTACACTACAGCATTTGGCT 1657
QY 301 -----AspProIleHisGluProProSerLeuProProGluProLysThrThrLysLeu 318
DB 1658 CCGGTCTCTCCAGCTAGCCCTCTGGAGTCTT-----GAGCCTTAAGCAGACGCGCTT 1711
QY 319 ---GlyLnaArgArgGlnSerSerArgProValLysProProLysAspValProAsp 337
DB 1712 CCCCCTATGCGTGAAGAGAGTGGTGGCCCATCAAGCCCAAGCAAGACTTGCTGAC 1771
QY 338 SerGlnGlnHisProAlaProGluLysSerLysValSerGlnGlnLeuLysCysCys 357
DB 1772 TCTCAGACACACACACAGAGCTCTTAAGAAAGAAAGCTTTCAGAACGTTAAACAATTCG 1831
QY 358 SerGlyTlleuLysGlnMetPheAlaLysLysHisAlaAlaTyrAlaTyrProPheTyr 377
DB 1832 AATGGCATTTTGAAGAGTACTCTCTTAAGAACATGCTCTCATGCTTGCTTCTAT 1891
QY 378 LysProValAspValGluAlaLeuGlyLeuHisAspTyrCysAspIleIleLysHisPro 397
DB 1892 AAACCGATGATCTTCTGACATTTGGCTGACATGACTACATGACATTAAGACACCC 1951
QY 398 MetAspMetSerThrTlleLysSerLysLeuGlnAlaArgGlnTyrArgAspAlaGlnGln 417
DB 1952 ATGAGACTGACACTGCTCAAGCGGAAGATGAGAACCGTATTACCGGATGACAGGAG 2011
QY 418 PheGlyAlaAspValArgLeuMetPheSerAsnCysTyrLysTyrAsnProProAspHis 437
DB 2012 TTTCGCTGATGATAGCGCTTATGTTCTCCAACTGCTATTAAGTAACTCCCAAGATCAC 2071
QY 438 GluValValAlaMetLysArgLysLeuGlnAspValPheGlnMetArgPheAlaLysMet 457
DB 2072 GATGTTGTGGCAATGGACAAAGCTACAGAGATGATTGAGTTCCGTTATGCGCAAGATG 2131
QY 458 ProAspGlnProGluGlnProValValAlaValSerSerProAlaValProProPro--- 476
DB 2132 CCGAGTGAACCACTAAGAACAGGCGCTTTACAGTCTCTACTGCAATGCCCTCGGCTTG 2191
QY 477 ThrLysValValAlaProProSerSerSerSerSerSerSerSerSerSerSerSer--- 494
DB 2192 GCCAAATCGCTTTCAGAGTCTCCAGTGAAGAAAGTACAGTACAGCTCTCTGAGGAA 2251
QY 495 -----AspSerAspSerSerThrAspAspSerGln 504
DB 2252 GACGAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2311
QY 505 GlnGluArgAlaGlnArgLeuAlaGlnLeuGlnGlnGlnLeuLysAlaValHisGlnGln 524
DB 2312 GAAGAAAGGCGCTATGCTTACAGCAACCTACAGAAACAGCTTCGGGCAAGTACATGAACAA 2371
QY 525 LeuAlaAlaLeuSerGlnProGlnGlnAsnLysProLysLysLysGluLysAspLysLys 544
DB 2372 CTGGCTGCTCTGTCCAGGCTCCAAATATCCAAAGCCCAAGAGAGAGAGAGAGAGAGAGAG 2431
QY 545 GluLysLys 547
DB 2432 AAAAAAAA 2440

```

```

RESULT 9
BQ219777 893 bp mRNA linear EST 02-MAY-2002
LOCUS BQ219777
DEFINITION AGENCOURT 7644864 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6020711
5', mRNA sequence.
ACCESSION BQ219777
VERSION BQ219777.1 GI:20401177
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 893)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1325 row: d column: 24
High quality sequence stop: 675.
FEATURES
source location/Qualifiers
1..893
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6020711"
/clone_1ib="NIH MGC 70"
/tissue_type="epitheloid carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: Nct1;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 259 a 298 c 199 g 137 t
ORIGIN
Alignment Scores:
Pred. No.: 1,24e-74 Length: 893
Score: 1436.00 Matches: 276
Percent Similarity: 96.22% Conservative: 4
Best Local Similarity: 94.85% Mismatches: 6
Query Match: 39.46% Indels: 5
DB: Gaps: 2
US-09-700-590A-22 (1-688) x BQ219777 (1-893)
QY 63 LeuGlnTyrLeuLeuArgValValLeuLysThrLeuTyrLysHisGlnPheAlaTyrPro 82
DB 3 CTGCAATACCTGCTGAGATGGTGCACAGACACTATGAGAAACACAGTTTGATGGCT 62
QY 83 PheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIleIleLys 102
DB 63 TTCACAGAGCTGTGATGCGCTCAAGCTGAACCTCCGATTTACTATAGATCATTTAA 122
QY 103 ThrProMetAspMetGlyThrTlleLysLysArgLeuGlnAsnAsnTyrTyrTrpAsnAla 122
DB 123 AGCCCTATGATATGGGAACATTAAGAAAGCGCTTGAAAAACACTATTACTGGAATGCT 182
QY 123 GlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrTlleTyrAsnLysPro 142
DB 183 CAGGAATGATCCAGACACTTCAACATATGTTTACCAATTTTACATCTACAAAGCGCT 242
QY 143 GlyAspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLysIleAsn 162
DB 243 GAGATGACATATGCTTAAATGGAGAGAGCTCTGAGAAAGCTCTTCTTGCAAAATTAAT 302
QY 163 GluLeuProThrGluGlnThrGluIleMetIleValGlnAlaLysGlyArgGly 182

```


Db 796 GCTCCCAAGCCCAAGAGACCCCAACCATCATTCGCGGCGCACCCCAAGGCTG 855
 Qy 283 -----LysThrLysLysGlyValLysArgLysAlaAspThrThr 296
 Db 856 GGGAGAGACCAAAAAAGGAGAGTGGAAAAAGGAAAAAGCCCAAAATCCAAACCAATCC 915
 Qy 296 rProthr 298
 Db 916 CCCCACC 922
 RESULT 11
 LOCUS B0646784 951 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8298482 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6270294
 5', mRNA sequence.
 ACCESSION B0646784
 VERSION B0646784.1 GI:21770956
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 951)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM2445 row: d column: 07
 High quality sequence stop: 636.
 Location/Qualifiers
 1. 951
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6270294"
 /clone_1ib="NIH_MGC_100"
 /issue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pORF7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH-MGC
 library."
 BASE COUNT 271 a 305 c 226 g 148 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,54e-73 Length: 951
 Score: 1408.00 Matches: 271
 Percent Similarity: 94.83% Conservative: 4
 Best Local Similarity: 93.45% Mismatches: 11
 Query Match: 36.69% Indels: 4
 DB: 14 Gaps: 0
 US-09-700-590A-22 (1-688) x B0646784 (1-951)
 Qy 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 Db 89 ATGTCTCGGAGAGAGCGCCCTGGGACAGATTGGAATCTGCCAGTAATGGGGGATGGA 148
 Qy 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 149 CTAGAAACTTCCCAATGTCTACACACAGGCCCAAGCCCAACCCCAAGCAGCAACGCA 208
 Qy 41 AlaSerThrAsnProProProGluThrThrSerAsnProAsnLysProLysArgGlnThr 60
 Db 209 GCCAGACCAACCCCGCCCGCCAGAGACTCTCAACCTTAACAGGCCAAGAGGAGACC 268
 Qy 61 AsnGlnLeuGlnThrLeuLeuArgValValLeuLysThrLeuThrLysHisGlnPheAla 80
 Db 269 AACCACTGCATACCTGCTCAGAGGTGCTCAACACATATGGAACACCAAGGATTTGCA 328
 Qy 81 TTPProPheGlnGlnProValAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 Db 329 TGCCCTTTCAGCAGCGCTGTGATGCCGTCAACCTCAACCTCTATTACTTAAGATC 388
 Qy 101 LLeuLysThrProMetAspMetGlyThrLLeuLysArgLeuLysAsnTyrTyrTTP 120
 Db 389 ATTAACCGCTATGATATGGAATGGAACATTAAGAAAGCGTTGGAAAAACAATTAATCG 448
 Qy 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCystrIleTyrAsn 140
 Db 449 AATGCTCAGGAATGTATCCAGGACTTCAACACTATGTTTAAATTTGTTACATCTACAAC 508
 Qy 141 LysProGlyAspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLys 160
 Db 509 AAGCTCGAGATGACATGACTTAAATGAGCAAGACTCTGAAAAAGCTCTTTGCAAAAA 568
 Qy 161 LLeuAsnGluLeuProThrGluGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 Db 569 ATTAATGAGTACCTCCACAGAAAGAACCGAGATCATATATGTCAGCAAAAGAAAGAA 628
 Qy 181 ArgGlyArgLysGluThrGlyThrAlaLysProGlyValSerThrValProAsnThrThr 200
 Db 629 CTGGGAGAGAAAGAAACAGAGACAGCAAAACCTGGGTTTCCACGGTATCAACAAACAACT 688
 Qy 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsnPro-ProProValGlnAl 220
 Db 689 CAAGCATGCACTCTCCGAGACCCGAGACCCCTCCAGCAATCTCTCTGTGAGAGC 748
 Qy 220 aThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMetTh 240
 Db 749 CAGCCTCACCCCTTCTCCGTGCGCTACCCCGAGACTCATCTGTCACACCCCTCTCTTGA 808
 Qy 240 rValValProProGlnProLeuGln-ThrProProProVal-ProProGlnProGlnPro 259
 Db 809 CGTGGCTCCCTCCCAACCCACTGAGAAAGCCCGCCAGTGGGCCCCAGACCAACACC 868
 Qy 260 ProProAlaProAlaProGlnProValGlnSerHisProProIleIleAla-AlaThrTr 279
 Db 869 CCACCGCTTCAGGCTCCAGCCCGGAAAGAGCCACCCATTTTCGGGGGCCACCCC 928
 Qy 279 oglnProValLysThrLysLys 286
 Db 929 ACAGCCCGGAGAAACCAAGAA 950
 RESULT 12
 LOCUS B0645818 925 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8350724 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6287189
 5', mRNA sequence.
 ACCESSION B0645818
 VERSION B0645818.1 GI:21769990
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 925)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2489 row: d column: 06
 High quality sequence stop: 684.

FEATURES

source
 1. .925
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6287189"
 /clone_id="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB1; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 261 a 295 c 220 g 147 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 4.73e-70 Length: 925
 Score: 1357.50 Matches: 264
 Percent Similarity: 93.33% Conservative: 2
 Best Local Similarity: 92.63% Mismatches: 11
 Query Match: 37.30% Indels: 9
 DB: 14 Gaps: 3

US-09-700-590a-22 (1-688) x B0645818 (1-925)

Qy 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAnLeuProValMetGlyAspGly 20
 Db 89 ATGCTCGGAGAGAGGGCCCTGGACGAGATTGAGAAATCTCCAGTAAATGGGGATGA 148
 Qy 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaSerAla 40
 Db 149 CTAGAACTTCCCAATGATCTACACACAGGGCCGACCCCAAGCCAGCCCAACCA 208
 Qy 41 AlaSerThrAsnProProProProGluThrSerAsnProAsnLysProLysArgGlnThr 60
 Db 209 GCCAGCACCAACCCGCCGCCAGAGACCTTCAACCAACCCCAAGGCGAGACC 268
 Qy 61 AsnGlnLeuGlnIleuLeuArgValIleuLysThrLeuTrpLysHisGlnPheAla 80
 Db 269 AACCACTGCAATACCTCTGCTCAGAGTGGTCTCAAGACCTATGGAAACACAGTTTGA 328
 Qy 81 TrpProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 Db 329 TGGCTTTCCAGACACCTGTGGATGCCGTCAAGCTGAACTCTCCGATTAATAGATC 388
 Qy 101 IleLysThrProMetAspMetGlyThrIleLysLysValGluGlnAsnAsnTyrTyrTrp 120
 Db 389 ATTAAAGCGCTATGATATGGGAACAATAAGAAAGCGCTTGGAAAAACAATATTACTCG 448
 Qy 121 AsnAlaGlnGlnCysIleGlnAspPheAsnThrMetPheThrAsnCysTyrIleTyrAsn 140
 Db 449 AATGCTCGAATGATATCAGAGACTTCAACACTATGTTTACAAATGTTACATCTCAAC 508
 Qy 141 LysProGlyAspAspIleValLeuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLys 160
 Db 509 AAGCTGAGATGATCATGTCTTAATGGAGAAAGCTCTGGAAAAAAGCTCTTCTTGA 568
 Qy 161 IleAsnGlnLeuProThrGlnGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 Db 569 ATTAATGAGTATCCACAGAGAAACCGAGATCATGATATGTCAGGCAAAAGAGAGAGA 628

Qy 181 ArgGlyArgLysGlnThrGlyThrAlaLysProGlyValSerThrValProAsnThrThr 200
 Db 629 CGTGGAGGAAAGAAACAGGAGACGACAAACCTGGCTTCCACGATACCAACACAACT 688
 Qy 201 GlnAlaSerThrProProGlnGlnThrGlnThrProGlnProAsnProProValGlnAla 220
 Db 689 CAACATTCGATCTCTCCGACAGACCAACCCCTCAGCAATCTCTCTCTGGACAGCC 748
 Qy 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMetThr 240
 Db 749 ACGCTACACCTTCTCTGCGGTACACCCCGGACTTATCTCCAGACCCCTGTAGAGACA 808
 Qy 241 -ValValProProGlnProLeuGlnThrProPro--ProValProProGlnProGlnPro 259
 Db 809 AGTGTGCTCCCAACCCCTGCGAGACGCCCCCGGAGGGGCCCCCAACACCAACC 868
 Qy 259 GProProAlaProAlaProGlnProValGlnSerHisPro--ProIleIleAlaIleThr 278
 Db 869 C-----CCAAC-CCGTTTCAGGGTTCCTCCAGGCGGGATTAAGAAAGCCCC 912
 Qy 278 rProGlnProVal 282
 Db 913 CCCACCCCTTTA 925

RESULT 13

B0653187

LOCUS

AGENCOURT_8341467 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6267841

DEFINITION

5', mRNA sequence.

ACCESSION

B0653187

VERSION

B0653187.1 GI:2177359

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 975)

AUTHORS

NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-1@mail.nih.gov

Tissue Procurement: CGAP (Stanford)

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2438 row: n column: 02
 High quality sequence stop: 656.

FEATURES

source

1. .975

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6267841"

/clone_id="NIH_MGC_100"

/tissue_type="hepatocellular carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pOTB1; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 275 a 304 c 246 g 145 t 5 others

ORIGIN

Alignment Scores:

Pred. No.: 1.03e-69 Length: 975
 Score: 1352.00 Matches: 263
 Percent Similarity: 90.91% Conservative: 7
 Best Local Similarity: 88.55% Mismatches: 23
 Query Match: 37.15% Indels: 4
 DB: 14 Gaps: 2

US-09-700-590a-22 (1-688) x BQ653187 (1-975)

QY 1 MetSerzlaGusErgiYProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 DB 89 ATGCTCCGAGAGAGCGCCCTGGGACGAGATTGGAATCTGCCAGTAATGGGAGATGA 148
 QY 21 LeuGluThrSerGlnMetSerThrThGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 DB 149 CTGAAACCTTCCCAATATGTCTACAACAGGCGCCAGCCCAACCCAGCCAGCCAGCA 208
 QY 41 AlAserThrAsnProProProGlnThrSerAsnProAsnLysProLysArgGlnThr 60
 DB 209 GCCAGCACCAACCCCGCCCGCCAGAGACCTCCAACTTAACAGCCCAAGAGGCAACC 268
 QY 61 AsnGlnLeuGlnThrLeuLeuArgValValLeuThrLeuTrpLysHisGlnPheAla 80
 DB 269 AACCAATCGAATTCCTGCTCAGAGTGCTCTCAAGACATATGAAACACAGTTTGA 328
 QY 81 TrpProhGlnGlnProValAspAlaValLysLeuAsnLeuProAspTyrTyrLysIle 100
 DB 329 TGGCTTTCCAGCAGCGCTGTGATGCCGTCAAGCTCAAGCTCTGATTAATTAAGATC 388
 QY 101 IleLysThrProMetLaspMetGlyThrIleLysLysArgLeuGlnAsnGlyTyrTrp 120
 DB 389 ATTAACACGGCTTATGATATGGAAACATTAAGAGGCTTGGAAAACACTATTACTGG 448
 QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThraenCysTyrIleTyrAsn 140
 DB 449 AATGCTAGAGATATATCCAGACTTCAACACATATGTTTACAAATTTGTTACATTAAC 508
 QY 141 LysProGlyLysAspIleValIleuMetAlaGlnAlaLeuGlnLysLeuPheLeuGlnLys 160
 DB 509 AAGCTCGAGATGACATATGCTTAATGACAGAACTCTGGAAAACTCTTCTTCAAAA 568
 QY 161 IleAsnGluLeuProThrGlnGluThrGluIleuMetIleValGlnAlaLysGlyArgGly 180
 DB 569 ATTAATAGCTACCCACAGAAACGAGATCATGATAGCCAGGCAAAAGAAAGAGA 628
 QY 181 ArgGlyArgLysGluThrGlnThrAlaLysProGlyValSerThrValProAsnThrThr 200
 DB 629 CGTGGAGAAAGAAACAGGACAGCAAACTGGCGTTTCACGGTACCAACACAACT 688
 QY 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsnProProProValGlnAla 220
 DB 689 CAAGCATGCACTCTCCGAGAACCCAGACCCCTCAGCGAATCTCTCTGTCAGAGCC 748
 QY 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMet-Th 240
 DB 749 ACGCTCAGACCTTCTCCGCGTACCCCGGACCTCATCTGTCAGACCCCTGGCATGAC 808
 QY 240 rValValProGlnProLeu---GlnThrProProProValProGlnProGlnInr 259
 DB 809 AGTGGTGCCTCCGNNNNCACTGGCAAGGCCCGCCGAGGGGGCCCCCAAGCCACACC 868
 QY 259 oProPro---AlaProAlaProGlnProValGlnSerHisPro-ProIleIleAlaIat 278
 DB 869 CCCACCCCGGTCCAGGTCGCCAGCCCGGTAAGAGGACACCCACCATCATCGGGGGCC 928
 QY 278 hrProGlnProValLysThrLysLysGlyValLysArgLysAlaAsp 293
 DB 929 CCCCCCAGGCTGGGAGAGAAAAAAGGGGAGCGGAAAGGGAA 975

RESULT 14
 LOCUS BQ749323 815 bp mRNA linear EST 17-JUL-2002
 DEFINITION UI-M-FCO-byb-e-03-0-UI.r1 NIH_BMAP_FCO Mus musculus cDNA clone

IMAGE:5715818 5', mRNA sequence.
 BQ749323
 BQ749323.1 GI:21896110
 EST.
 VERSION BQ749323.1
 KEYWORDS house mouse.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 815)
 NIH-MGC <http://imgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/HLN at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pyx-5.
 Location/Qualifiers
 1. 815
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAG:5715818"
 /clone_1ib="NIH_BMAP_FCO"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pyx-acc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pyx-acc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 221 a 269 c 189 g 136 t

ORIGIN

Alignment Scores:
 Pred. No.: 7.16e-69 Length: 815
 Score: 1336.50 Matches: 252
 Percent Similarity: 94.10% Conservative: 3
 Best Local Similarity: 92.99% Mismatches: 15
 Query Match: 36.73% Indels: 2
 DB: 14 Gaps: 1

US-09-700-590a-22 (1-688) x BQ749323 (1-815)

QY 170 GluIleMetIleValGlnAlaLysGlyArgGlyArgLysGluThrGlnThrAla 189
 DB 2 GAGATCATGATATCTCCAGGCAAAAGAGAGCGAGGAGAAACAGGAGGAGCA 61
 QY 190 LysProGlyValSerThrValProAsnThrThGlnAlaSerThrProProGlnThrGln 209
 DB 62 AAGCTGTGTATCCACGGTACCAACACAACTCAAGCATCACTCTCCGAGACCCAG 121
 QY 210 ThrProGlnProAsn---ProProProValGlnAlaThrProHisProPheProAlaVal 228
 DB 122 ACGCTCAGCAACCTCTCTCAGCTGTGACGCGCAACATCACTCCCTTCTGCTGTC 181

QY 229 ThrProAspleuIleValGlnThrProValMetThrValProProGlnProLeuGln 248
 DB 182 ACCCGAGACCTCATTTGCCAGCCTCTGTCAAGCAATGGAGGCCCTCCAGCACTTCAAG 241
 QY 249 ThrProProProValProProGlnProGlnProProProAlaProAlaProGlnProVal 268
 DB 242 ACTCTTCACCGGTACCCCTCCAGCACCCCTCCAGCTGTCTCCAGCTGTGTG 301
 QY 269 GlnSerHisProProIleIleAlaAlaThrProGlnProValLysThrLysGlyVal 288
 DB 302 CAGAGTCACCCCGCATATATGCAACCCCTCCAGCTGTGAGACAAAGAGGAGTG 361
 QY 289 LysArgLysAlaAspThrThrThrProThrThrIleAspProIleHisGluProProSer 308
 DB 362 AAGAGGAAGAAGATACCAACCCCTTACCAACATCGACCCCATTCATGAGCCACCTCA 421
 QY 309 LeuProProGluProLysThrThrLysLeuGlyGlnArgGluSerSerThrProVal 328
 DB 422 CTGGCCCGAGCCCAAGACCGCCAAAGCTGGTCTCCGCGGAGAGCAGACGACTGTG 481
 QY 329 LysProProLysLysAspValProAspSerGlnGlnHisProAlaProGluLysSerSer 348
 DB 482 AAGCTCCAAAGAGATGACCGGACTCACAGACGCCAGAGGCGCCAGAGAGACAC 541
 QY 349 LysValSerGluGlnLeuLysCysCysSerGlyIleLeuLysGluMetPheAlaLysLys 368
 DB 542 AAGATCTCTGAGCACTAAAGTGGCGAGTGGCATCTCAAGAGATTTTGGCAAGAA 601
 QY 369 HisAlaAlaLysAlaLysProPheThrLysProValAspValGluAlaLeuGlyLysHis 388
 DB 602 CATGCTGCTATGCTGCTGCTTTCTACCAAGCTGTGAGTGTGAGGCACTGGTCTGAC 661
 QY 389 AspThrCysAspIleIleLysHisProMetAspMetSerThrIleLysSerLysLeuGlu 408
 DB 662 GACTACGTGACATCATCAATCAATCCATGACATGACATGACATCAATCAATCAATGAG 721
 QY 409 AlaArgGluThrArgAspAlaGlnGlnPheGlyAlaAspValArgLeuMetPheSerAsn 428
 DB 722 TCCGAGAGTACAGAGATGCCAGGAATTTGGTGTGATGTCCGATTATGTCTCAAC 781
 QY 429 CysThrLysThrAsnProProAspHisGluVal 439
 DB 782 TGCTACAGATACAC-CCCCCTGACCATGAGTG 813
 RESULT 15
 LOCUS B0650970 926 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT 848559 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6296078
 ACCESSION B0650970
 VERSION B0650970.1 GI:21775142
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 926)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LICM2503 row: f column: 15
 High quality sequence stop: 595.
 Location/Qualifiers

source 1..926
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6296078"
 /clone_lib="NIH_MGC_100"
 /issue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Size-selected >500bp for average insert size
 1 kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."
 BASE COUNT 267 a 290 c 218 g 151 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7..73e-68 Length: 926
 Score: 1319.50 Matches: 256
 Percent Similarity: 92.45% Conservative: 1
 Best Local Similarity: 92.09% Mismatches: 16
 Query Match: 36.26% Indels: 6
 DB: 14 Gaps: 2
 US-09-700-590a-22 (1-688) x B0650970 (1-926)
 QY 1 MetSerAlaGluSerGlyProGlyThrArgLeuArgAsnLeuProValMetGlyAspGly 20
 DB 89 ATGTCTCGGAGACCGGCTGGAGACAGATGTGAATCTGCCAGTAAGGGGAGTGA 148
 QY 21 LeuGluThrSerGlnMetSerThrThrGlnAlaGlnAlaGlnProGlnProAlaAsnAla 40
 DB 149 CTAGAACTTCCAAATGTCTTCAACACAGGCCCAAGCCCAAGCCAGCAAGCA 208
 QY 41 AlaSerThrAsnProProProProGluThrSerAsnProAsnLysProLysArgGlnThr 60
 DB 209 GCCAGCAACCAACCCCGCCGCCAGAGACTCCCAACCTTAACCAAGCCCAAGAGGCAACC 268
 QY 61 AsnGlnLeuGlnThrLysLeuArgValValLeuLysThrLeuTrpLysHisGlnPheAla 80
 DB 269 AACCAACTGCAATATCCGCTCAGAGTGTCTCAAGCACTATGAGAAACACAGATTGCA 328
 QY 81 TrpProPheGlnGlnProValAspAlaValLysLeuAsnLeuProAspThrLysLysIle 100
 DB 329 TGCCCTTTCCAGCAGCTGTGAGTCCGTCAAGCTGAACCTCCTGATTCTATAAGATC 388
 QY 101 IleLysThrProMetAspMetGlyThrIleLysLysArgLeuGluAsnAsnThrLysTrp 120
 DB 389 ATTAAAGCGCTTATGATATGAGAAACAATAAGAAAGCGCTTGAGAAACAATTACTCG 448
 QY 121 AsnAlaGlnGluCysIleGlnAspPheAsnThrMetPheThrAsnCysThrIleTyrAsn 140
 DB 449 AATCTCAGGAATATATCCAGGACTTCAACACTATGTTTCAAAATTTTTCATCTACAC 508
 QY 141 LysProGluAspAspIleValLeuMetAlaGluAlaLeuGluLysLeuPheLeuGlnLys 160
 DB 509 AAGCTGGAGATGACATAGTCTTAAATGACAGACCTCTGAGAAAGCTCTTCTGCAAAA 568
 QY 161 IleAsnGluLeuProThrArgLysGluThrGluIleMetIleValGlnAlaLysGlyArgGly 180
 DB 569 ATAATATAGCATACCAACAGAAAGAAACGAGATCATGATGTCAGGCAAAAGAGAGGA 628
 QY 181 ArgGlyArgLysGluThrArgLysThrAlaLysProGlyValSerThrValProAsnThrThr 200
 DB 629 CGTGGAGAGAAAGAAAGAGGACGCAAAACCTGGCTTTTACAGGTATCCAAACAACT 688
 QY 201 GlnAlaSerThrProProGlnThrGlnThrProGlnProAsnProProProValGlnAla 220
 DB 689 CAAGCATGACTCCCTCGACAGACCAAGCCCTCAGCGAATCTCTCTGTGAGAGGC 748

```
Qy 221 ThrProHisProPheProAlaValThrProAspLeuIleValGlnThrProValMet-Th 240
Db 749 ACSCCTCACCCCTTCTCTGCGTCACCCCGACCTCATGTCAGAACCCCTGCATGAC 808
Qy 240 rValValProProGlnProLeuGlnThr---ProProProValProProGlnProGlnPr 259
Db 809 AGTGTGCTCTCCACACCCCATGAAAAAGGCCCCCGCCAGAGCCCCCCCAGGTCAAAA 868
Qy 259 oProProAlaPro-----AlaProGlnProValGlnSerHisProPro 273
Db 869 ACCCCCAAC-CCGTTCAAGTTCCCCAGGCCCGGAACAGACCAACCCCCC 919
```

Search completed: March 14, 2003, 22:42:08
Job time : 2141 secs

THIS PAGE BLANK (USPTO)